1998

Modeling the distribution of woodpecker species in the Jura France and in Switzerland using atlas data

Claudine Tobalske
The University of Montana

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

MODELING THE DISTRIBUTION OF WOODPECKER SPECIES IN THE JURA, FRANCE, AND IN SWITZERLAND, USING ATLAS DATA

by

Claudine Tobalske

Diplôme d'Etudes Universitaires Générales, Université de Besançon, France, 1988

Maîtrise Biophysiologie Appliquée aux Productions Végétales, Université d'Angers, France, 1990

Diplôme d'Etudes Supérieures Spécialisées Ressources Naturelles et Environnement, Université de Nancy 1, France, 1991

Presented in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

University of Montana

Spring 1998

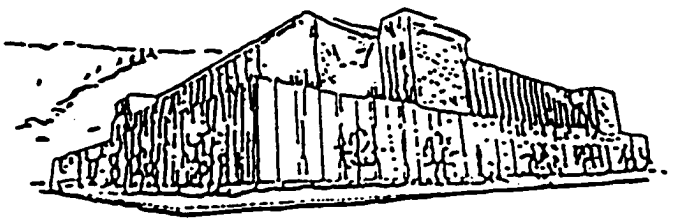
Approved by

Chairman, Board of Examiners

Dean, Graduate School

5-18-98

Date



Maureen and Mike
MANSFIELD LIBRARY

The University of **MONTANA**

Tobalske, Claudine, Ph.D., Spring 1998 Wildlife Biology

Modeling the distribution of woodpecker species in the Jura, France, and in Switzerland, using atlas data (137 pp.)

Director: Roland L. Redmond RLR

ABSTRACT

Wildlife-habitat relationship (WHR) models, that relate the distribution of species to characteristics of their environment, have become a tool commonly used in wildlife management and conservation. To assess the usefulness and identify some of the limitations associated with using breeding bird atlas data to generate such models, I used maps from a local atlas and extracted simple, broad-scale habitat variables using a GIS to develop logistic regression models for seven woodpecker species in the Jura, France: the Black (*Dryocopus martius*), Green (*Picus viridis*), Grey-headed (*P. canus*), Great Spotted (*Dendrocopos major*), Middle Spotted (*D. medius*), and Lesser Spotted (*D. minor*) woodpeckers, and the Wryneck (*Jynx torquilla*). Although the selected independent variables were in general agreement with known ecological requirements of the species, the reliability of the models was species-specific. Despite limitations inherent to working in a grid format, the models could be used to identify areas with high probability of woodpecker presence as targets for future censuses and fine-scale habitat studies. However, an independent validation performed by applying the models to two sites in Switzerland showed that model performance outside of the Jura was low. This poor predictive generality was probably the result of important differences between the variables used for model development, and those used for the validation, although additional factors may have contributed to model failure. To further explore this issue of model generality, new logistic models were developed for the green woodpecker using the atlas distribution maps from the two Swiss sites. The effect of scale on predictive accuracy was assessed by enlarging the mesh size of the breeding bird atlas maps and increasing the minimum mapping unit of the habitat map. Cross-area predictions were not successful, possibly because the variables used in the models did not reflect differences in landscape configuration between the sites. Model composition, classification results, and predictive generality were all affected by scale. This variability suggests that caution should be used when applying a model developed in one part of a species' range to another, and that particular attention should be given to the choice of the scale of analysis.

PREFACE

Models, because they represent a simplified view of reality, can help us understand natural patterns and processes, and thereby provide useful tools for the protection and management of our environment. The need to manage wildlife habitat has been recognized for many years (e.g., Leopold 1933). However, unprecedented rates of species extinctions in recent decades, caused by fragmentation and loss of habitat worldwide, have made conservation measures all the more urgent. In this context, the advent of computer technology over the past decade has led to the rapid growth of wildlife-habitat relationship (WHR) models (Verner *et al.* 1986). WHR models are developed for two main reasons: 1) to improve our understanding of the relationships between wildlife species and their habitat (i.e., to identify which factors affect the distribution and abundance of species; and 2) to predict the distribution and/or abundance of species (Morrison *et al.* 1992). In both cases, the ultimate goal of the model usually is related to management and/or protection of species and their habitat.

WHR models can be categorized as either deductive or inductive. The deductive approach relies on previous knowledge of species-habitat associations to derive model rules, and the actual species' distributions are not required. Examples of deductive models include Gap Analysis, the broad-scale mapping of potential habitat of vertebrate species across the United States (Scott *et al.* 1993); Habitat Evaluation Procedures, that assess environmental conditions at the species level (Flood *et al.* 1977); and Habitat Suitability Index models, that denote habitat suitability of a species as the geometric mean of environmental variables (Schamberger *et al.* 1982). With inductive models, habitat variables are gathered at or around known species' locations, and some type of statistical analysis is used to select which environmental parameters best account for any observed variation in the distribution and abundance of the species. Multivariable

statistical approaches, such as multiple regression, principal component analysis, discriminant function analysis, and logistic regression are commonly used to develop inductive WHR models (Capen 1981).

Inductive models are most useful to help formalize our understanding of a species' relation to its habitat, but they require that the species' distributions must be known. Because of the difficulty of obtaining distribution data over broad areas, habitat selection studies have often been conducted at a relatively fine scale, for example by collecting environmental data within close proximity of a reproduction site. Broad-scale distribution data, such as the outline of a species' range, can be used to develop WHR models, but these may be too crude to be used for habitat management. Another type of broad-scale distribution data is becoming available, and with a level of detail far superior to that of traditional range maps. Distribution atlases, in the form of systematically sampled grids, are now common, thanks to the coordinated efforts of dedicated volunteers. The size of the cell varies, usually from 1 or 2 km for local or regional studies (e.g. Glayre and Magnenat 1984, Thomas and Abery 1995) to as large as 30 km for entire countries (e.g. Robertson *et al.*1994). An increasing number of studies are looking at the potential of such atlases to provide information other than distribution *per se*, including prediction of species distribution from incomplete atlas data (Osborne and Tigar 1992), study of seasonal patterns of migration (Underhill *et al.*1992), estimation of population size (Robertson *et al.*1994, Heikkinen 1998), study of abundance and/or distribution changes (Thomas and Abery 1995, Bircham and Jordan 1996, Böhning-Gease and Bauer 1996), or impact of commercial afforestation on bird species diversity (Allan *et al.* 1997).

There are, however, limitations inherent to using atlas data. Robertson (1994) questioned the quality of the South African Bird Atlas because rare and inconspicuous

species were under-represented, and unequal sampling effort among grid cells may have introduced additional bias. Heikkinen (1997) also expressed concerns about the suitability of mesoscale atlas data to model the distribution of rare vascular plant species in Finland. Thomas and Abery (1995) found that the scale of the atlas cell size exerted a strong influence on estimates of butterfly species decline; population losses were underestimated in 10-km grid cells compared to 2-km grid cells. Finally, other users of gridded data (which need not be presented in an atlas format) have raised additional concerns, such as the presence of spatial autocorrelation (Smith 1994), or the failure of WHR models due to the positioning of the map grid (Fielding and Haworth 1995).

The overall goal of this dissertation is to explore the question: What is the potential for atlas data to be used as the dependent variable to develop broad-scale WHR models? Broad-scale habitat variables can be easily compiled with Geographic Information Systems and associated with other digital products, such as classified satellite images or digital elevation models (DEMs), to provide a source of independent variables for many areas. The real question, though, is not whether atlas data can be used to develop WHR models -- they can; but whether the resulting models meet their intended purpose. Few studies have used atlas data for WHR model development. Osborne and Tigar (1992) used the survey results of 55 quarter-degree grid cells (about 24 km by 27.5 km) of the Lesotho breeding bird atlas to develop logistic regression models for three bird species. Errors of commission (predicted probability of species occurrence in cells were no bird had been recorded during the field survey), though not insignificant, were not thought to limit the applicability of the models. On the contrary, the authors concluded that the model outputs could be more reliable than the atlas data, especially for rare and inconspicuous species likely to be missed during field surveys. Gates *et al.*(1995) used a similar approach with data from the breeding bird atlas of

England (10 km by 10 km grid) to model the distribution of eight farmland bird species that showed signs of decline over the past 25 years. Although their models were "able to predict with a good degree of accuracy the present patterns of abundance of the eight species", they did not accurately predict the populations in 1969. Parker (1996) used the one-eighth degree grids from the Swaziland bird atlas to gain insight into the nature and relative strengths of relationships between the distribution of bird species and environmental variables. Although his models successfully identified explanatory variables for 335 bird species, the grid-based approach complicated the interpretation of the results.

In this dissertation I aim to provide additional information about the utility and limitations of atlas data for the study of wildlife-habitat associations through the specific example of woodpecker species in three study sites in France and Switzerland. Each of the following chapters addresses a specific issue: the ability of the models to predict species' current distribution (Chapter 1); the ability of models to predict species' distribution under conditions different from those used to develop the models (i.e, model universality; Chapter 2); and the sensitivity of both classification success and model universality to scale (Chapter 3).

REFERENCES

Allan, D. G., J. A. Harrison, R. A. Navarro, B. W. Van Wilgen, and M. W. Thompson. 1997. The impact of commercial afforestation on bird populations in Mpumalanga province, South Africa - insights from bird-atlas data. Biol. Conserv. **79**:173-185.

Bircham, P. M. M, and W. J. Jordan. 1996. A consideration of some of the changes in distribution of "common birds" as revealed by *The New Atlas of Breeding Birds in Britain and Ireland.* Ibis **139**:183-186.

Böhning-Gease, K., and H. G. Bauer. 1996. Changes in species abundance, distribution, and diversity in a Central European bird community. Conserv. Biol. **10**, 175-187.

Capen, D.E., ed. 1981. The use of multivariate statistics in studies of wildlife habitat. USDA Fro. Serv. Gen. Tech. Rep. RM-87. Rocky Mtn. Forest and Range Exp. Sta., Fort Collins, CO.

Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conserv. Biol. **9**, 1466-1481.

Flood, B. S., M. E. Sangster, R. S. Sparrow, and T. S. Baskett. 1977. A handbook for habitat evaluation procedure. USDI Fish and Wildl. Serv. Res. Publ. No. 132.

Gates, S., D. W. Gibbons, P. C. Lack, and R. J. Fuller. 1993. Declining farmland bird species: modelling geographical patterns of abundance in Britain. Pages 153-177 *in* P. J. Edwards, R. M. May, and N. R. Webb, eds. Large-Scale Ecology and Conservation Biology. Blackwell Sci. Publ., London.

Glayre, D., and D. Magnenat. 1984. Oiseaux Nicheurs de la Haute Vallée de l'Orbe. Nos Oiseaux 398. Fascicule spécial du volume 37.

Heikkinen, R. K. 1998. Can species patterns of rarities be predicted from mesoscale atlas data? A case study of vascular plants in the Kevo reserve. Biol. Conserv. **83**, 133-143.

Leopold, A. 1933. Game management. Charles Scribner's Sons, New York, NY.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1992. Wildlife-Habitat Relationships. Univ. of Wisconsin Press, Madison.

Osborne, P. E., and B. J. Tigar. 1992. Interpreting bird atlas data using logistic models: an example from Lesotho, Southern Africa. J. Appl. Ecol. **29**, 55- 62.

Parker, V. 1996. Modeling the distribution of bird species in Swaziland in relation to environmental variables. Ostrich **67**, 105-110.

Robertson, A., R. E. Simmons, A. M. Jarvis, and C. J. Brown. 1994. Can bird atlas data be used to estimate population size? A case study using Namibian endemics. Biol. Conserv. **71**, 87-95.

Schamberger, M., A. H. Farmer, and J. W. Terrel. 1982. Habitat suitability index models: an introduction. USDI Fish and Wildl. Serv. FWS/OBS-82/10.

Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Butterfield, S. Caicco, C. Groves, T. Edwards, Jr., J. Ulliman, H. Anderson, F. D'Herchia, and R. G. Wright. 1993. Gap analysis: a geographic approach to protection of biodiversity. Wildl. Monogr. 123.

Smith, P. A. 1994. Autocorrelation in logistic regression modelling of species' distribution. Global Ecol. And Biogeography Letters **4**, 47-61.

Thomas, C. D., and J. C. G. Abery. 1995. Estimating rates of butterfly decline from distribution maps: the effect of scale. Biol. Conserv. **73**, 59-65.

Underhill, L. G., R. P. Prys-Jones, J. A. Harrison, and P. Martinez. 1992. Seasonal patterns of occurence of Palaearctic migrants in southern Africa using atlas data. Ibis **134** (Suppl.1), 99-108.

ACKNOWLEDGEMENTS

I am grateful to A. Jovéniaux, P. Géroudet, D. Glayre, and the other authors of the three breeding bird atlases for allowing me to use their distribution data. I also wish to thank the Parc Naturel Régional du Haut-Jura (Lajoux, France) and the Wildlife Spatial Analysis Lab (WSAL; University of Montana, Missoula, USA) for providing access to a GIS station. The Laboratoire Environnement et Paysages of the Geography Department of the University of Besançon, France, provided the raw Landsat TM image required to run the MERGE program. C. Winne and T. Taddy (WSAL) helped me download the image and run the MERGE program; I am thankful to all the members of WSAL for sharing their technical knowledge and for providing a wonderful work environment. R. Redmond, my advisor, and D. Hutto, P. Landres, D. Pletscher, and H. Zuuring, the members of my dissertation committee, provided helpful suggestions during the study, and useful comments on earlier drafts of the dissertation.

Special thanks to B. Tobalske for his constant support and encouragements, and for sharing with me his passion for woodpeckers.

This work is dedicated to

Bret and Brendan Tobalske

for reminding me that even when models fail, life goes on!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 Page

CHAPTER 2

CHAPTER 3

APPENDIX 1

OBJECTIVES

CHAPTER 1

1. To assess whether habitat variables extracted at the 575-ha scale for seven woodpecker species in the Jura *département* , France, are suitable for broad-scale distribution modeling;
2. To determine if the presence and absence of woodpeckers in the Jura *département* can be predicted using simple, broad-scale variables, and to assess whether the level of detail of the variables, the number of variables entering the models, and the ecology of the species affects the ability of the models to correctly predict presences and absences;
3. To assess whether models derived from the Jura breeding bird atlas could be useful for the broad-scale management and conservation of woodpecker species in the Jura *département* .

CHAPTER 2

1. To test the ability of the models developed in Chapter 1 to predict the distribution of woodpecker species in a location, and with a data set, different from those used for model development.

CHAPTER 3

1. To determine if the associations between the green woodpecker's presence and simple, broad-scale habitat variables are strong and consistent enough to allow model universality (i.e., the ability of a model to predict species distribution under conditions other than those used for model development);
2. To assess the effect of changing the scale of distribution data (atlas cell size) and that of habitat variables (MMU) on model classification results and model

universality.

CHAPTER 1

MODELING THE DISTRIBUTION OF WOODPECKER SPECIES IN THE JURA, FRANCE, USING ATLAS DATA

Key words: breeding bird atlas, France, grid data, habitat model, logistic regression, scale, woodpecker.

ABSTRACT

Populations of several woodpecker species have declined in Europe. Although local-scale habitat requirements of woodpeckers have received much attention, broad-scale habitat relationships remain poorly studied for many species. To develop broad-scale models that may be of use to managers of wildlife habitat, I used data from a local breeding bird atlas and landscape variables obtained from a geographic information system (GIS) to study the distribution of seven woodpecker species in the Jura, France: the black (*Dryocopus martius*), green (*Picus viridis*), grey-headed (*P. canus*), great spotted (*Dendrocopos major*), middle spotted (*D. medius*), and lesser spotted (*D. minor*) woodpeckers, and the Wryneck (*Jynx torquilla*). I created three types of bird-habitat models using logistic regression to assess whether including more independent variables and using different forest maps influenced the predictive ability of the models. Univariate comparisons of presences and absences revealed that the selected independent variables, although crude, were in general agreement with known ecological requirements of the species. Modeling reliability was more influenced by the species' ecology than by the type of model used. Best results were obtained for those species that behaved as specialists with regards to the habitat variables used in the

analysis, or for which broad-scale habitat variables could be used as surrogates for finer-scale habitat requirements, such as the black and middle spotted woodpeckers. The models, despite some limitations inherent to working in a grid format, could be used to identify areas with high probability of woodpecker presence as a target for future censuses and fine-scale habitat studies.

INTRODUCTION

The decline of many European bird species over the past decade is a source of concern among biologists. Major shifts in human activities, such as intensification of agriculture (Bobbink and Willems 1993) or afforestation of previously open grounds (Moss *et al.* 1979), are responsible for a rapid modification of the traditional, semi-natural landscapes. Although the decline appears to be more pronounced for grassland species (Gibbons *et al.* 1993, Gates *et al.* 1994, Böhning-Gaese and Bauer 1996), concerns have risen for forest species as well; for example, the French federal agency responsible for forest management was recently mandated to protect biodiversity in publicly owned forests (ONF 1993). Woodpeckers are particularly at risk because of their strong dependency on forest characteristics such as snags and old growth, that conflict with traditional silvicultural management (Angelstam and Mikusinski 1994). Indeed, woodpecker populations have shown signs of decline in many European countries (Glutz von Blotzheim and Bauer 1980, Cramp 1985, Pettersson 1985, Taianen 1985, Wesolowski and Tomialojc 1986, Mikusinski and Angelstam 1998).

Seven woodpecker species can be found in the Jura *département*, France: the black woodpecker, green woodpecker, grey-headed woodpecker, great spotted woodpecker, middle spotted woodpecker, lesser spotted woodpecker, and Wryneck.

The Three-toed woodpecker (*Picoides tridactylus*) is also known to occur, although sightings are rare. The seven common species occupy a broad range of habitats and vary in their ecological requirements from more open-landscape species such as the green woodpecker and the Wryneck, to forest species such as the middle spotted woodpecker. Population declines have been documented in the Jura; three species, the grey-headed and middle spotted woodpecker, and the Wryneck, have either reached a critical level or shown a severe decline in the last 20 years (Jovéniaux 1993); in fact, the great spotted woodpecker is the only woodpecker that is not cited on the Jura Red List of Breeding Birds. The black woodpecker is listed locally as a species with very low numbers, but in France its range has been expanding over the past 20 years (Cuisin 1980). Perhaps less at risk than the other species, the great spotted and black woodpeckers are nonetheless of strong interest to conservationists: they are considered to be keystone species (Johnsson *et al.* 1990, Tjernberg *et al.* 1993) because of the large number of species using their abandoned cavities (Cuisin 1988, Johnsson *et al.* 1993). The black, grey-headed, and middle spotted woodpeckers are on the list of species 'afforded special protection' by the European Community Bird Directive (Ehrlich *et al.* 1994).

Fine-scale habitat features, such as snags, are certainly crucial to the maintenance of woodpecker populations; studies conducted at a broader scale, however, suggest that landscape characteristics should be considered as well (Angelstam 1990, Wiktander *et al.* 1992, Tjernberg *et al.* 1993). Until recently, two factors complicated such landscape-scale studies: (1) the difficulty of analyzing large quantities of spatial data; and (2) the paucity of census data over large areas. These limitations can now be largely overcome. Geographic information systems (GIS), computer-based systems for the manipulation and analysis of spatially-distributed data (Johnson 1990), have

revolutionized the analysis of habitat data. Variables can be easily extracted over broad tracts of land, either by classifying satellite images, or by digitizing existing maps, which can then be overlaid on species distribution maps in the GIS to extract habitat relationships. Availability of census data has also improved: distribution atlases, in the form of systematically sampled grids, are becoming more common and are usually available for broad areas. The size of the cell varies, usually from one or two km for regional studies (e.g. Glayre and Magnenat 1984, Thomas and Abery 1995) to as large as 30 km for entire countries (e.g. Robertson *et al.* 1994). An increasing number of researchers are looking at the potential of such atlases to provide information other than distribution *per se*, including prediction of species distribution from incomplete atlas data (Osborne and Tigar 1992), study of seasonal patterns of migration (Underhill *et al.*1992), estimation of population size (Robertson *et al.*1994, Heikkinen 1998), study of distribution changes (Thomas and Abery 1995, Bircham and Jordan 1996), or impact of commercial afforestation on bird species diversity (Allan *et al.* 1997). Using species distribution maps to obtain broad-scale correlations between species and their environment appears to be another potential use of atlases, especially if habitat data can be entered into a GIS. I addressed this possibility using a specific example.

A breeding bird atlas was recently released for the Jura *département*, using 575-ha cells (2.3 by 2.5 km) (Jovéniaux 1993). Eight years of censuses provide a clear picture of the distribution of the seven woodpecker species listed above, but no quantitative study of woodpecker-habitat relationships at the landscape scale has been conducted. A good knowledge of bird-habitat associations at several geographic scales is the key to the protection of woodpeckers and their habitat; such knowledge, in the Jura, exists mainly at the local scale, and with more detail for some species than others. No comprehensive study of woodpecker-habitat relationships has been conducted for

the whole *département.* The availability of a GIS prompted me to use bird distribution data from the atlas to address the following questions:

1. How do habitat variables extracted at the 575-ha scale for woodpecker species in the Jura *département* compare to the existing scientific and anecdotal knowledge of habitat associations for these species? Are such variables suitable for broad-scale distribution modeling?

2. To what extent can the presence and absence of woodpeckers in the Jura *département* (obtained from atlas data) be predicted using simple, broad-scale variables? Does the level of detail of the variables, and the number of variables entering the models, affect modeling results? Does the ecology of the species influence the ability of the models to correctly predict presence and absence?

3. How useful would woodpecker-habitat models derived from the Jura breeding bird atlas be for the broad-scale management and conservation of these species in the Jura *département* ?

Beyond the particular case of woodpeckers in the Jura *département*, I wished to identify some of the uses and limitations of information derived from bird distribution atlases.

STUDY AREA

The Jura *département* (hereafter referred to as Jura) is a 5055 km^2 administrative entity located in the eastern part of France, along the Swiss border (Fig. 1). Named after the mountain range that runs along its eastern side, it can be divided into several physiographic regions which follow an altitudinal gradient from northeast to

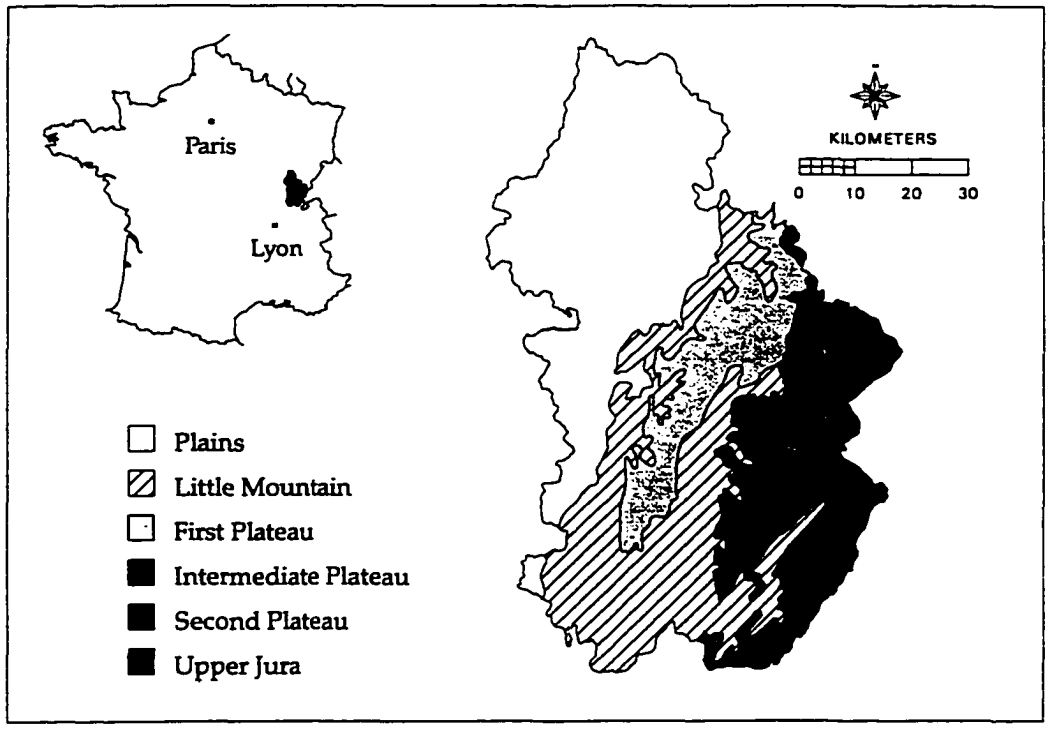


Figure 1. Location of the study area in France, and physiographic entities of the Jura (adapted from IFN 1993).

southwest: the Plains, Little Mountain, First, Intermediate, and Second Plateaux, and Upper Jura regions (Fig. 1). This gradient of increasing elevation influences climate, vegetation, and patterns of human land-use. With 45% forest cover (Jovéniaux 1993), the Jura is one of the most forested *départements* in France (Fig. 2). Highly managed deciduous forests (rich in oak, *Quercus* spp.) dominate areas below 450 m and, as elevation increases, are replaced with beech (*Fagus sylvatica*) and mixed forests (beech and white fir, *Abies alba).* Norway spruce (*Picea abies*) predominates from 800 m to 1300 m. Several large forested areas of the First Plateau that were once deciduous were largely converted to monocultures of fir following World War II (e.g., Moidons and Poligny forests; Fig. 2). High elevation coniferous forests dominated by Norway spruce grow in the wettest and coldest portions of the Jura; these uneven-aged stands are managed through selective cutting. Large tracts of agricultural crop lands are located in the Plains; elsewhere agricultural tracts are small and devoted essentially to cattle grazing. The Jura is relatively undeveloped compared to other French *départements.*

METHODS

Digital database

Presence/absence data

Bird censuses were conducted yearly from 1985 to 1992 (inclusive) by members of the Groupe Ornithologique du Jura (full details available in Jovéniaux 1993). The sampling grid was created by dividing 20 tiles of the French 1:50,000 scale Lambert grid into 96, 575-ha cells, which resulted in 1920 cells, and sampling only the 945 cells that fell inside of the Jura boundary. I recreated the census grid using the GIS

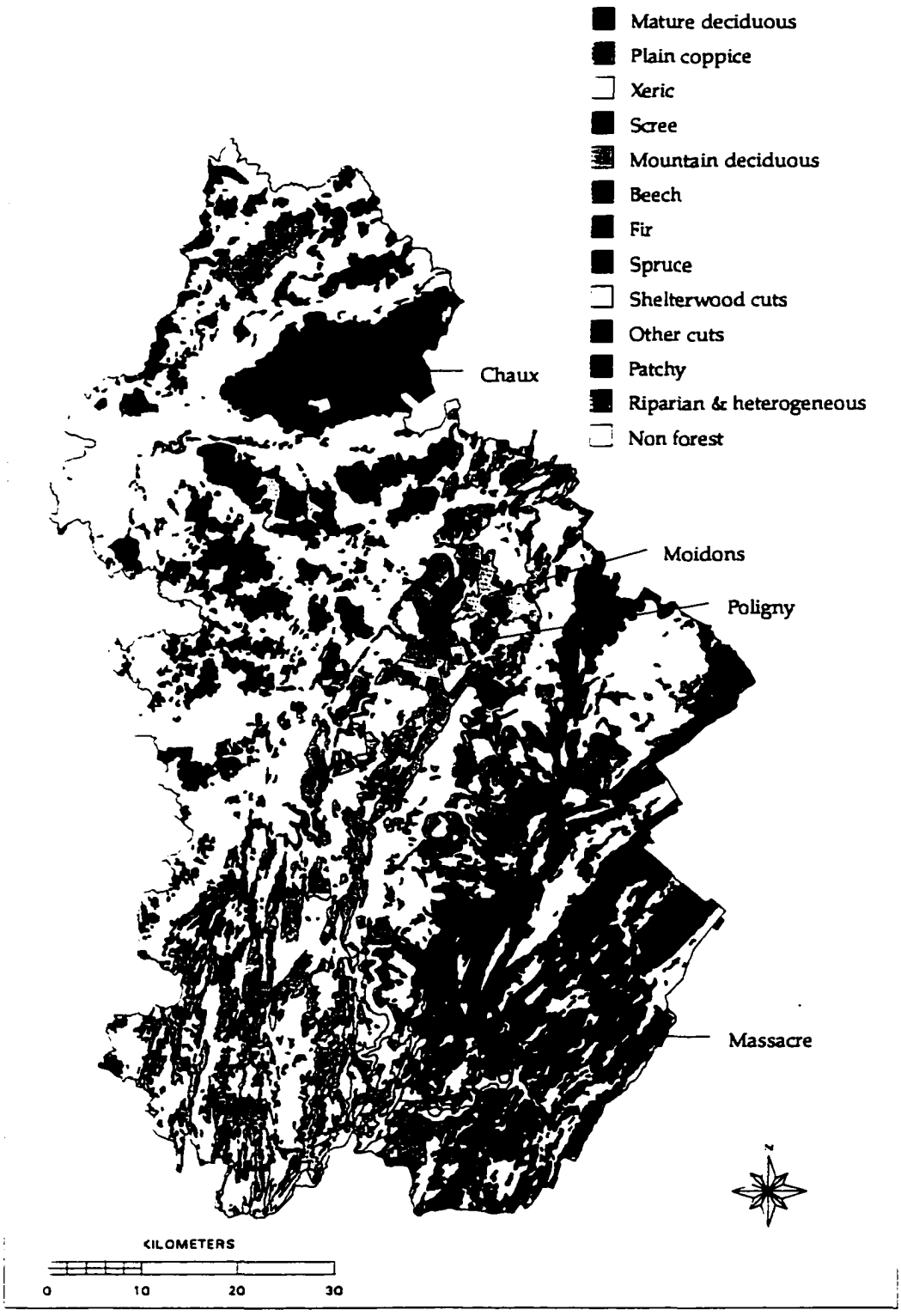


Figure 2. Forest cover map of the Jura, France. Major forests referred to in the text are identified (adapted from IFN 1993).

software PC Arc/Info 3.4.2. (ESRI 1994). Because habitat data were lacking for parts of grid cells that fell outside the Jura, I eliminated from the analysis only those cells that were less than two-thirds inside the study area, assuming that vegetation data obtained for 66% of a cell were representative of the entire composition of that cell. Eighty-nine of the 945 cells were dropped, leaving 856 cells for analysis.

Jovéniaux (1993) conducted a thorough examination of the results was conducted at the end of each field season, to insure that all cells were censused with the same intensity regardless of topography, and to limit the observer bias that could be introduced by a pre-acquired knowledge of the avifauna of certain cells.

Presence/absence data for model development were obtained from the atlas as follows. If a species's breeding status in a cell was listed as either "possible", "probable", or "certain", the species was considered to be present in that cell. This approach was preferred over other groupings (such as considering only "certain" cells as presence cells) because even a "possible" breeding status required that a singing male, or mating calls, had been heard within suitable habitat during the breeding period (Schifferli *et al.* 1980). I estimated this definition to be sufficiently different from absent (no singing bird censused in the cell during the breeding season) to justify calling it a presence. Also, locating active nests (the condition required for a cell to receive a status of "certain") can be difficult for woodpecker species in the Jura (pers. obs.). Indeed, the percent of presence cells that received a status of "certain" ranged from 0% for the grey-headed woodpecker, to only 16% for the wryneck (Jovéniaux 1993). Using only those cells classified as "certain" in the analysis would have underestimated the abundance of breeding cells, and the resulting sample size would have been too small for analysis for most species (maximum 41 cells for the great spotted woodpecker).

Habitat variables

I selected 17 variables (13 forest cover classes, mean elevation, average slope, cover class richness, and edge density) based on data availability and on their suspected importance to woodpeckers as assessed from literature and personal knowledge of the species' habitat requirements.

Forest cover

A forest cover map, recently released for the whole Jura at a 1:200,000 scale, was purchased from the Inventaire Forestier National (IFN). Although this was a paper map, it was not transferred onto a stable medium such as mylar prior to digitizing because of its thickness (reduced sensitivity to moisture changes) and the absence of folds. I digitized the map using PC Arc/Info v. 3.4.2. (ESRI 1994). Polygons were coded according to the following thirteen-class legend (IFN 1980; variable name in parentheses): mature deciduous forests (MatDecid), where mature, large-diameter trees (mostly oak *Quercus* sp.) compose more than 50% of the canopy; plain coppice (Coppice), also dominated by oak, but with a lower proportion of large-diameter trees and a canopy cover less than 50%; xeric forests (Xeric), characterized by small-diameter trees and often a dense understory of boxwood (*Buxus sempervirens*); scree forests (Scree), restricted to humid scree slopes; mountain deciduous forests (MtnDecid), rich in oak and beech of medium-size diameter, often with a dense coppice of beech; beech forests (Beech), pure or mixed with conifers; fir forests (Fir), dominated by conifers and composed of at least 75% of white fir; spruce forests (Spruce), dominated by conifers and composed of at least 75% of Norway spruce; shelterwood cuts (ShelterCut), where conifers (usually fir) are planted under an overstory of mature deciduous trees eventually removed; other cuts (OtherCut), i.e., any

other type of regeneration cut (e.g., clearcuts, or plantations in open fields); patchy forests (Patchy), i.e., small woodlots and clumps of trees of varied composition (although predominantly deciduous), used for wood production by local people; riparian and heterogeneous forests (RipHetero), usually small clumps of trees connected by a network of edges or wooded pastures; and non forested (Open). A simplified version of the forest map was created by aggregating the above thirteen classes into five, more 'general' classes (Table 1, Fig. 2), to compare modeling results obtained from the full and the simplified maps.

Elevation and slope

A map of elevation was produced by manually overlaying a grid of points printed on an acetate sheet on top of 1:100,000 topographic maps and recording elevation at each point. The grid was created by systematically locating 16 points within each atlas cell, which roughly corresponded to one point every 600 m. Elevation points were entered into the GIS database, and assigned the corresponding x-y coordinates (variable Meanelev). A digital elevation model was then created using the software PC SEM (ESRI 1994) and used to derive a slope layer (variable Slope).

Richness and edge density

Forest richness (the number of different forest classes; variable Richness) was computed for each cell as a measure of habitat fragmentation. Edge density (variable Edgeden) was obtained after combining all the forest classes into one (edge being the limit between forest and non-forest) for two reasons. First, edge density computed among the different forest types was strongly correlated with richness, and thus redundant. Second, woodpecker species such as the green woodpecker are associated

Table 1. Aggregation of 13 forest cover classes from the Inventaire Forestier National (IFN) map of the Jura, France, into five, more general classes.

IFN forest cover class (variable name)	Broad cover type class
Mature deciduous forests (MatDecid)	Deciduous
Plain coppice (Coppice)	Deciduous
Xeric forests (Xeric)	Deciduous
Scree forests (Scree)	Deciduous
Mountain deciduous forests (MtnDecid)	Deciduous
Beech forests (Beech)	Mixed
Fir forests (Fir)	Mixed
Spruce forests (Spruce)	Mixed
Shelterwood cuts (ShelterCut)	Planted
Other regeneration cuts (OtherCut)	Planted
Patchy forests (Patchy)	Patchy
Riparian and heterogeneous forests (RipHetero)	Patchy
Not forested (Open)	Open

with edges between forest and open areas (Short 1982, Clergeau and Chefson 1988). The GIS was used to compute, for each cell, the total length of the line(s) representing the boundary between forest and non-forest. This value (in meters) was divided by the area of the cell (in hectares) to obtain edge density (m/ha).

Statistical analyses

I used PC Arc/Info to overlay the IFN forest cover map with the distribution grids of the seven woodpecker species and to obtain composition (in percent of each forest class), edge density, and forest richness, for each atlas cell. Mean elevation and slope were entered directly into a spreadsheet. All statistical analyses were performed with SPSS v. 4.0.4. (SPSS Inc. 1990).

Because of their presentation (in a grid format), data obtained from atlases may be prone to spatial autocorrelation, a phenomenon frequently encountered when a variable is mapped onto a geographic space (Sokal and Oden 1978). Most classical statistics assume independence of the observations, so statistical problems are to be expected; for example, spatial autocorrelation can lead to a poor model fit. I computed Moran's I, Geary's c, and conducted a semivariance analysis to assess the amount of spatial autocorrelation in the data (Appendix 1).

To identify bird-habitat relationships at a broad scale, I conducted univariate analyses for the 17 variables by comparing values for these variables between presence and absence cells using a Mann-Whitney U-test. This non-parametric test was selected because most variables departed strongly from normality. For each species, I retained the variables exhibiting a statistically significant difference in mean rank between presence and absence cells ($P \leq 0.01$), then reviewed existing literature on woodpecker-habitat relationships (anecdotal, and detailed, quantitative studies) for comparison with

the results obtained from the test.

I used logistic regression (LR) to create models to classify presence and absence data. This multivariable statistical approach has recently come into favor in studies of wildlife-habitat relationships (Morrison *et al.* 1992) because it permits the prediction of binary attributes such as presence/absence (McCullagh and Nelder 1983). LR is particularly suited to highly skewed data that include many zeros (Stowe *et al.* 1993, Green *et al.* 1994), which was the case for the forest cover data.

Even if, for most species, the number of absences largely outnumbered that of presences, I was reluctant to eliminate cells from an analysis. A higher number of absences can be an advantage in presence-absence studies, because such cells are expected to show more variability (Capen *et al.* 1986, Pereira and Itami 1991). I computed univariate LR models for all 17 variables for each bird species, then forced all of them into one model; sign reversal, compared to the univariate LR, indicated that variable reduction was required. Rather than use stepwise selection procedures which can lead to biologically implausible models (Hosmer and Lemeshow 1989), I applied four steps recommended by Hosmer and Lemeshow (1989): 1) variables were selected based on their statistical significance in univariate LR analysis (using the Wald statistics); 2) one of a pair of strongly correlated variables was eliminated; 3) perfectly discriminated variables were excluded; and 4) any variable judged to be ecologically important, even if non statistically significant in the univariate LR analysis, was kept. A reduced set of variables was obtained for each species.

Cut-off points were selected to maximize both the number of presences and absences correctly classified, although preference was given to presences. When categorizing the probabilistic output of a logistic model, a cut-off of 0 classifies all cases as presences, whereas a cut-off of 1 classifies all cases as absences. I used a histogram

of estimated probabilities to roughly locate the zone of cut-off values over which a reversal from more presences correctly classified, to more absences correctly classified, was observed. I then varied the cut-off value by increments of 0.01 until reversal was obtained. I retained as a cut-off the value just before the reversal, thereby always obtaining slightly more presences than absences correctly classified. For example, if the reversal occurred between 0.36 and 0.37, then 0.36 was the selected cut-off.

Because there is no way to unequivocally obtain a 'best' model, I developed several models per species in the following fashion. I divided the retained independent variables into subsets of decreasing size (from all the retained variables, to each variable considered singly) and developed models from these subsets. When one or more variable in a model showed an R value (partial correlation between the dependent and the independent variables) smaller than 0.01, the model was eliminated from further analysis, unless I estimated that this (these) variable(s) were important to the species, based on known habitat requirements. Three types of models were created for each species. For 'biological' models, I retained the models that contained the maximum number of statistically significant variables ($R \geq 0.01$), plus ecologically meaningful variables. The same rule applied to 'universal' models, but those were built from the simplified forest cover map. Finally, 'parsimonious' models had to include fewer variables than the 'biological' ones, with a maximum of three, and the regression coefficients for all independent variables had to be statistically significant ($R \geq 0.01$).

In certain cases, two or more possible models occurred for each model type (as would for example be the case for the parsimonious type, if various combinations of three variables all resulted in statistically significant R values). A 'best model' (one per model type) was then selected based on the percent presences and absences correctly classified, and visual comparison of prediction maps. Prediction maps were obtained

by recoding, in the GIS, each cell with the corresponding probability of presence $P = 1/(1 + e^{-Z})$, where Z represented the linear function of independent variables. At this point, three models (one per model type) were retained for each species.

I checked linearity in the logit for each model by regressing the logit of the dependent variable on each independent variable. For variables exhibiting a non-linear relationship, logistic regression models were created that included squared terms, a common transformation procedure because most species show unimodal responses to environmental gradients (Gauch and Chase 1974). Each squared term was first added individually to the basic model, then combinations of squared terms were added, and finally all the squared terms were added together to the equation. The behavior of the parameters of the model was examined to decide whether or not squared terms should be kept. Squared terms were kept if the R statistics of the independent variables remained significant (at the 0.01 level), and if there was no sign reversal of the original variables. Percent presences and absences correctly classified by these new models, and the stability of their variables' coefficients (based on a ten-fold validation procedure; see below), were compared to those of the basic model. Models with squared terms were preferred over basic models only if they improved the percent correctly classified of both presences and absences (even by a few percent), while producing models at least as stable as the basic ones. Three final, 'full' models, were retained for each species. I did not try to add interaction terms because none appeared biologically meaningful.

Validation is a critical step of model development, because it determines how much confidence can be placed in the models (Morrison *et al.* 1992). I conducted a cross-validation (Capen *et al.* 1986, Livingston *et al.* 1990) by randomly assigning the

cells to ten groups, creating a model using nine of them, and testing it on the excluded group. The process was repeated nine times, each group being excluded in turn. Percent presence and absence correctly classified, as well as the value of the variable coefficients (β) were obtained for the ten subsets and compared to the parameters of the full model. A coefficient of variation (CV) was calculated to assess the stability of each variable (Capen *et al.* 1986). For negative β, the absolute value of the CV was used. When a large CV (i.e., at least three times as large as that of the other variables) was observed (a sign of instability), the model was re-run after eliminating the corresponding problem variable. The validation test was carried out again on the resulting model. This procedure was repeated until all problem variables were eliminated. This resulted in three stable models for each species.

The fit of each model to the logistic curve was assessed by conducting a chi-square test as described in Loftsgaarden *et al.* (1992). Model probabilities were ranked, split into ten groups of almost equal size (four groups of 85, six of 86), and summed within each group to obtain predicted presences; this value was subtracted from 1 to obtain predicted absences. Observed values were simply the number of true presences and absences within each group. The measure of fit was obtained by conducting a chi-square test on the 20 cells; a good fit of the model to the logistic curve should result in a small chi-square value, i.e., a large P value.

Because high classification rates can be obtained by chance when number of presences and absences are unequal (Morrison 1969, Capen *et al.* 1986), I computed percent correct classification due to chance alone and estimated the improvement of the models over chance using Cohen's Kappa (K) statistic (Titus *et al.* 1984): $K = (P_o - P_c) / (1 - P_c)$, where P_o = overall percent classification correct after modeling and P_c =

percent classification correct due to chance alone (detailed formulas in Appendix 2). I computed $Z = K / SE_k$ and obtained the corresponding P-value. A high P-value (e.g., $P > 0.05$) indicates a Kappa not statistically different from zero, i.e., the model did not bring any improvement over chance classification.

RESULTS

Univariate analysis

The number of statistically significant independent variables ranged from (at the 0.001 level) four for the Wryneck, to as many as eleven for the black and green woodpeckers (Table 2). Elevation was strongly significant for every species, and may be used to discriminate between the black woodpecker (positive correlation) and the other species (negative correlation). Despite a mean cell elevation of 792 m, however, the black woodpecker was also found breeding in cells with mean elevation as low as 190 m and as high as 1350 m (Table 3). Such a broad amplitude was found for the other species as well, except for the middle spotted woodpecker (range 205-316 m) which appeared to be the only true lowland species. Associations with slope were generally similar in magnitude and direction as with elevation, which is not surprising considering how strongly correlated these two variables were (Spearman's $r = 0.68$, $P \leq 0.0005$; Table 4).

The cover types most strongly associated with the presence of each woodpecker species could be identified from differences in mean rank between presence and absence cells (Table 2). For the black woodpecker for example, cells with high percentages of fir and spruce forests were more likely to be coded as presence than cells

Table 2. Mann-Whitney U-tests (* = $P \leq 0.01$, ** = $P \leq 0.001$) of presences and absences (N) of seven woodpecker species with 17 independent variables in the Jura, France.

	Black Woodpecker		Green Woodpecker		Grey-headed Woodpecker		Great spotted Woodpecker		Middle spotted Woodpecker		Lesser spotted Woodpecker		Wryneck	
Independent variable	Absent (736)	Present (120)	Absent (443)	Present (413)	Absent (820)	Present (36)	Absent (326)	Present (530)	Absent (821)	Present (35)	Absent (7170)	Present (139)	Absent (787)	Present (69)
Meanelev	393	646**	511	340**	440	175**	470	403**	439	189**	465	241**	437	332**
Slope	403	587**	449	406*	440	166**	468	405**	439	180**	452	306**	427	448
Edgeden	428	429	397	462**	432	343	439	422	436	260**	435	397	427	449
Richness	428	431	418	440	434	310*	429	428	435	268**	439	374*	432	392
MatDecid	441	355**	402	457**	418	666**	376	461**	413	796**	401	570**	430	416
Coppice	438	373**	406	453**	421	610**	392	451**	422	584**	412	516**	430	419
Xeric	426	442	416	442	432	349*	440	422	432	349*	434	400	428	433
Scree	433	404	412	447**	430	386	428	429	430	386	424	451	425	467*
MtnDecid	438	370**	407	452**	434	313**	438	423	435	284**	441	363**	427	446
Beech	405	572**	468	386**	430	385	421	433	428	444	437	383*	434	361*
Fir	400	602**	483	370**	433	330*	446	418	434	307**	448	328**	437	335**
Spruce	410	541**	475	379**	431	364*	452	414**	431	364*	441	366**	433	375*
ShelterCut	439	367**	414	444	429	407	417	436	432	358	431	415	426	458
OtherCut	426	443	439	418	429	413	400	446*	430	396	431	414	432	387
Patchy	441	352**	398	462**	432	345	449	416	434	311*	425	446	419	535**
RipHetero	433	404	412	447	434	302**	452	414*	436	262**	442	359**	430	410
Open	464	212**	376	485**	432	353	482	396**	436	247**	422	460	414	589**

Table 3. Means and standard deviations (in parentheses) along with ranges of 17 variables for the presence of seven woodpecker species in the Jura, France.

	Black Woodpecker		Green Woodpecker		Grey-headed Woodpecker		Great spotted Woodpecker		Middle spotted Woodpecker		Lesser spotted Woodpecker		Wryneck	
Variable	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Meanelev	792 (295)	190 - 1350	405 (199)	188 - 1236	254 (107)	188 - 696	485 (273)	190 - 1287	240 (25)	205 - 316	305 (133)	190 - 834	376 (149)	190 - 797
Slope	12 (7)	1 - 37	8 (6)	0 - 37	3 (2)	0 -11	8 (6)	0 - 37	3 (2)	1 - 10	6 (5)	0 - 19	9 (6)	0 - 24
Edgeden	14 (8)	0 - 35	15 (8)	0 - 42	11 (8)	0 - 31	14 (8)	0 - 37	9 (9)	0 - 31	13 (8)	0 - 31	15 (8)	0 - 34
Richness	4.7 (1.6)	2 - 9	5 (2)	1 - 10	4 (1)	1 - 6	5 (2)	1 - 9	4 (1)	1 - 6	4 (2)	1 - 8	4 (2)	1 - 10
MatDecid	5.0 (17)	0 - 96	8 (17)	0 - 97	32 (30)	0 - 94	11 (22)	0 - 100	55 (31)	3 - 100	19 (26)	0 - 100	3 (9)	0 - 46
Coppice	1 (4)	0 - 33	2 (5)	0 - 44	6 (10)	0 - 44	2 (6)	0 - 55	5 (8)	0 - 33	3 (7)	0 - 44	1 (3)	0 - 16
Xeric	5 (15)	0 - 74	3.1 (10)	0 - 74	absent	absent	2 (9)	0 - 63	absent	absent	2 (9)	0 - 60	2 (6)	0 - 43
Scree	0 (1)	0 - 13	1 (3)	0 - 25	absent	absent	1 (3)	0 - 25	absent	absent	1 (3)	0 - 25	1 (3)	0 - 14
MtnDecid	3 (9)	0 - 49	8 (13)	0 - 79	4 (14)	0 - 67	7 (14)	0 - 79	1 (6)	0 - 33	4 (9)	0 - 47	7 (11)	0 - 45
Beech	8 (13)	0 - 57	2 (5)	0 - 42	2 (4)	0 - 18	4 (10)	0 - 64	2 (5)	0 - 15	2 (7)	0 - 55	0 (1)	0 - 8
Fir	23 (28)	0 - 92	2 (7)	0 - 64	1 (8)	0 - 50	7 (18)	0 - 92	absent	absent	1 (9)	0 - 91	1 (3)	0 - 21
Spruce	15 (26)	0 - 98	1 (5)	0 - 69	absent	absent	3 (13)	0 - 98	absent	absent	0 (0)	0 - 1	0 (0)	0 - 2
ShelterCut	1 (7)	0 - 78	3 (9)	0 - 83	4 (16)	0 - 85	3 (10)	0 - 85	0 (2)	0 - 12	2 (9)	0 - 85	2 (5)	0 - 31
OtherCut	3 (5)	0 - 20	3 (6)	0 - 47	3 (5)	0 - 22	4 (6)	0 - 47	2 (4)	0 - 13	4 (7)	0 - 47	2 (4)	0 - 18
Patchy	2 (3)	0 - 15	3 (4)	0 - 25	2 (3)	0 - 11	3 (4)	0 - 25	1 (3)	0 - 11	3 (5)	0 - 25	6 (7)	0 - 25
RipHetero	3 (6)	0 - 37	4 (7)	0 - 37	1 (5)	0 - 26	3 (7)	0 - 67	0 (1)	0 - 6	3 (7)	0 - 37	3 (6)	0 - 26
Open	28 (19)	0 - 82	59 (25)	0 - 100	44 (28)	0 - 96	59 (28)	0 - 100	30 (30)	0 - 86	56 (28)	0 - 100	71 (19)	18 - 100

Table 4. Spearman rank correlations among 17 habitat variables in the Jura, France ($N = 856$; * = $P \leq 0.0005$, two-tailed test).

	Meanelev	Slope	Edgeden	Richness	Mat-Decid	Coppice	Xeric	Scree	Mtn-Decid	Beech	Fir	Spruce	Shelter-Cut	Other-Cut	Patchy	Rip-Hetero	Open
Meanelev	1.00																
Slope	0.68*	1.00															
Edgeden	0.22*	0.45*	1.00														
Richness	0.23*	0.43*	0.44*	1.00													
MatDecid	-0.65*	-0.58*	-0.20*	-0.15*	1.00												
Coppice	-0.54*	-0.44*	-0.06	0.04	0.72*	1.00											
Xeric	0.13*	0.50*	0.36*	0.40*	-0.28*	-0.22*	1.00										
Scree	0.03	0.28*	0.15*	0.29*	-0.19*	-0.13*	0.27*	1.00									
MtnDecid	0.18*	0.41*	0.32*	0.54*	-0.39*	-0.29*	0.45*	0.22*	1.00								
Beech	0.46*	0.35*	0.03	0.27*	-0.15*	-0.17*	-0.05	-0.07	-0.05	1.00							
Fir	0.59*	0.36*	0.06	0.28*	-0.39*	-0.32*	0.01	-0.07	0.02	0.36*	1.00						
Spruce	0.58*	0.35*	0.10	0.01	-0.26*	-0.22*	-0.05	-0.09	-0.18*	0.42*	0.25*	1.00					
ShelterCut	0.02	0.00	0.08	0.37*	-0.09	-0.05	0.01	0.16*	0.36*	-0.10	-0.06	-0.21*	1.00				
OtherCut	0.08	0.03	0.05	0.44*	0.01	0.06	-0.01	0.06	0.24*	0.03	0.09	-0.13*	0.27*	1.00			
Patchy	-0.08	0.12*	0.35*	0.34*	0.02	0.13*	0.02	0.08	0.04	-0.03	-0.04	-0.04	-0.03	-0.06	1.00		
RipHetero	0.23*	0.30*	0.36*	0.39*	-0.41*	-0.32*	0.33*	0.09	0.36*	-0.05	0.16*	-0.12*	0.05	0.07	-0.02	1.00	
Open	-0.41	-0.34	-0.10	-0.31	-0.02	0.04	-0.16	0.01	-0.22	-0.44	-0.35	-0.27	-0.12	-0.32	0.20	-0.01	1.00

dominated by shelterwood cuts and open areas. In contrast, the green woodpecker and the Wryneck were rarely present in cells that had a high percentage of coniferous forests, but were often associated with cells that had a high proportion of open habitats. Presences of these two species were also associated with patchy forests. The grey-headed, great spotted, middle spotted and lesser spotted woodpeckers were present in cells where mature deciduous forest was proportionally important. The middle spottedwoodpecker was never associated with four forest types (Xeric, Scree, Fir and Spruce) and the grey-headed woodpecker was never associated with three of them (Xeric, Scree, and Spruce). Differences in edge density were only significant for two species: mean rank was statistically higher for presences than for absences for the green woodpecker, but the contrary was observed for the middle spotted woodpecker ($P \leq 0.001$ in both cases). Presences for this species also ranked statistically lower than absences for cover type richness.

Logistic models

Several independent variables never entered any of the biological and parsimonious models (Table 5). I systematically eliminated Slope from the pool of variables used to build the models whenever Meanelev was also significant. The high correlation between these two variables made it necessary to retain only one of them, and elevation is more meaningful in the ecology of the species. Richness, Xeric, Scree, and RipHetero were included in the pool of variables for certain species, but always came out to be non significant ($R < 0.01$). They were not 'forced in' because I did not consider them to be crucially important for any species, based on known ecological requirements. ShelterCut and OtherCut entered the biological model of the Black Woodpecker, but a high coefficient of variation was obtained for OtherCut during the

validation process (|CV| = 174.9). Removing the variable greatly increased the overall stability of the model. A high CV was obtained for ShelterCut as well (|CV| = 66.87), but the CV of the other variables remained reasonable, so this variable was kept. Such a high CV was an exception though, and CVs of most variables entering the final models were low (< 20%), indicating model stability.

The number of independent variables in the biological models varied from three for the grey-headed and great spotted woodpeckers, to six for the black woodpecker; for the parsimonious models they varied from one for the middle spotted woodpecker to four for the black and green woodpeckers; and for the biological models, from two for the grey-headed, great and lesser spotted woodpeckers, to four for the green woodpecker (Table 5). The addition of squared terms was significant ($P \leq 0.01$) and improved the models' performances of only three models (black woodpecker, parsimonious; and green woodpecker, parsimonious and universal; Table 5).

Model performances differed among species and among the three types of models, but all the models classified presences and absences better than chance alone (Table 6). Improvement over chance classification ranged from 39.4% for the biological model of the black woodpecker, to only 12.1% for the parsimonious model of the Wryneck, which was still a statistically significant improvement ($P \leq 0.001$). When improvement over chance was not considered, best prediction rates were achieved for the black, middle spotted, and grey-headed woodpeckers, for all three model types ($P_o > 75\%$); however, correcting for chance showed that models for the black, middle spotted, and green woodpeckers were most successful in correctly classifying presences and absences ($K > 30\%$; Table 6). If a mean of the K values of the three models is computed for each species, improvement over chance ranks as follow: black (38.13%),

Table 5. Linear function of independent variables for 21 logistic models of three types (B = biological, P = parsimonious, U = universal) for seven woodpecker species in the Jura, France.

Black Woodpecker
B: - 1.1626 - (0.0203*MtnDecid) + (0.0161*Beech) + (0.0329*Fir) + (0.033*Spruce) - (0.0142*ShelterCut) - (0.0277*Open)
P: - 1.4824 + (0.0866*Fir) - (0.0007*Fir^2) + (0.0389*Spruce) - (0.0279*Open)
U: - 1.5132 + (0.0332*Mixed) - (0.0241*Open) - (0.011*Planted)
Green Woodpecker
B: - 1.7149 + (0.0556*Edgeden) + (0.0135*MatDecid) - (0.0643*Spruce) + (0.0179*Open)
P: - 2.4416 + (0.034*Edgeden) + (0.0191*MatDecid) + (0.0508*Open) - (0.0003*Open^2)
U: - 3.3055 + (0.0257*Edgeden) + (0.0293*Decid) + (0.0615*Open) - (0.0003*Open^2)
Grey-headed Woodpecker
B: - 1.3833 - (0.0066*Meanelev) + (0.0212*MatDecid) + (0.0346*Coppice)
P: - 3.8975 + (0.0337*MatDecid) + (0.0653*Coppice)
U: - 0.9522 - (0.0085*Meanelev) + (0.0229*Decid)
Great spotted Woodpecker
B: 0.852 + (0.0274*MatDecid) + (0.0397*Coppice) - (0.0106*Open)
P: 2.714 - (0.0019*Meanelev) - (0.0223*Open)
U: 0.6713 + (0.0149*Decid) - (0.0084*Open)
Middle spotted Woodpecker
B: - 1.6992 - (0.0085*Meanelev) + (0.0543*MatDecid) - (0.01076*OtherCut)
P: - 4.6207 + (0.0617*MatDecid)
U: -1.1739 - (0.0123*Meanelev) + (0.0562*Decid) - (1.094*Planted)
Lesser spotted Woodpecker
B: - 1.5965 + (0.0208*MatDecid) + (0.0238*Coppice) - (0.021*MtnDecid) - (0.0239*Beech) - (0.0437*Fir)
P: - 2.0048 + (0.0257*MatDecid) + (0.0385*Coppice)
U: - 1.6139 + (0.0112*Decid) - (0.0386*Mixed)
Wryneck
B: - 3.9794 - (0.083*Fir) - (0.1957*Spruce) + (0.1259*Patchy) + (0.0211*Open)
P: - 4.8079 + (0.128*Patchy) + (0.0301*Open)
U: - 3.9871 + (0.0532*Patchy) - (0.0792*Mixed) + (0.0232*Open)

middle spotted (33.17%), green (30.97%), lesser spotted (20.97%), great spotted (19.70%), grey-headed (15.50%), and Wryneck (13.03%). This ranking order would be the same if only the best model was considered for each species.

Highest values of P_o corresponded to highest values of K for six out of seven species. For the grey-headed woodpecker, the biological and universal models had very similar P_o values (76.8% and 76.5%), but improvement over chance was higher for the universal model (17.6% vs 15.8%). Biological models provided the highest improvement over chance for the black and middle spotted woodpeckers and for the Wryneck; universal models worked best for the green and grey-headed woodpeckers; and parsimonious obtained highest K for the great and lesser spotted woodpeckers (Table 6). A graphical display of the best models' predictions (Fig. 3) permits a visual comparison of predicted vs. true presences.

Fit of the models to the logistic curve varied among species. Smallest departures were obtained for the green, middle spotted, and lesser spotted woodpeckers, and for the Wryneck (chi-square test, $P \geq 0.05$, df = 8; Table 7). In all cases, the pattern of distribution of the residuals was random. Large chi-square values were sometimes obtained, and always resulted from a very large value in only one of the ten intervals. For the biological model of the grey-headed woodpecker, for example, presences in one interval had a small predicted value (0.22) but an observed value of 2. Even though two presences in a group of 85 is very small, this alone resulted in a chi-square value of 14.19, causing the overall chi-square value to be high ($X^2 = 20.41$, Table 7). Because differences between observed and expected values were small for the other intervals, I feel that the model did fit the logistic curve reasonably well.

Table 6. Regression equation statistics (standard error of the estimate *SEE*, coefficient of determination R^2) and classification results of three types of logistic models (B = biological, P = parsimonious, U = universal) for seven woodpecker species in the Jura, France. Computation formulas for P_o, P_c, K, SE_k and Z are presented in Appendix 2.

	Black Woodpecker			Green Woodpecker			Grey-headed Woodpecker			Great spotted Woodpecker		
	B	P	U	B	P	U	B	P	U	B	P	U
SEE	0.298	0.295	0.299	0.466	0.476	0.466	0.191	0.193	0.190	0.470	0.466	0.474
R^2	0.270	0.280	0.260	0.135	0.096	0.135	0.102	0.081	0.106	0.068	0.082	0.048
P_o	0.780	0.778	0.772	0.633	0.632	0.669	0.768	0.756	0.765	0.612	0.625	0.591
P_c	0.638	0.643	0.637	0.500	0.499	0.500	0.724	0.719	0.715	0.514	0.514	0.513
K	0.394	0.378	0.372	0.266	0.265	0.339	0.159	0.130	0.176	0.202	0.229	0.160
SE_k	0.045	0.046	0.045	0.034	0.034	0.034	0.055	0.055	0.054	0.035	0.035	0.035
Z	8.67	8.25	8.22	7.79	7.77	9.89	2.87	2.38	3.25	5.75	6.52	4.55
P	***	***	***	***	***	***	<0.0025	<0.01	<0.0025	***	***	***

	Middle spotted Woodpecker			Lesser spotted Woodpecker			Wryneck		
	B	P	U	B	P	U	B	P	U
SEE	0.164	0.168	0.164	0.353	0.357	0.359	0.256	0.258	0.264
R^2	0.318	0.284	0.315	0.091	0.064	0.054	0.117	0.102	0.065
P_o	0.898	0.875	0.872	0.629	0.750	0.601	0.681	0.668	0.674
P_c	0.841	0.818	0.814	0.528	0.648	0.543	0.629	0.623	0.625
K	0.360	0.315	0.308	0.214	0.291	0.125	0.140	0.121	0.130
SE_k	0.079	0.072	0.072	0.036	0.046	0.037	0.045	0.044	0.044
Z	4.58	4.35	4.30	5.91	6.28	3.35	3.13	2.76	2.94
P	***	***	***	***	***	***	<0.0025	<0.005	<0.0025

*** $P < 0.001$

Table 7. Chi-square tests of fit of logistic regression models to the logistic curve (* = $P \geq 0.05$, ** = $P \geq 0.1$) for three types of models (B = biological, P = parsimonious, U = universal) for seven woodpecker species in the Jura, France.

		Presences	Absences	Sum
Black woodpecker	B	14.65	1.62	16.27
	P	38.59	4.49	43.08
	U	12.74	2.48	15.22*
Green woodpecker	B	6.09	7.89	13.98*
	P	7.89	6.10	13.99*
	U	8.00	7.26	15.26*
Grey-headed woodpecker	B	20.01	0.40	20.41
	P	24.92	0.62	25.54
	U	33.11	0.38	33.49
Great spotted woodpecker	B	7.61	12.65	20.26
	P	2.65	3.39	6.04**
	U	7.70	8.41	16.11
Middle spotted woodpecker	B	4.87	0.63	5.50**
	P	14.98	0.22	15.10*
	U	0.71	0.01	0.72**
Lesser spotted woodpecker	B	12.19	2.23	14.42*
	P	32.47	5.11	37.58
	U	8.92	1.39	10.31**
Wryneck	B	7.88	0.68	8.56**
	P	12.67	1.33	14.00*
	U	11.47	1.84	13.31**

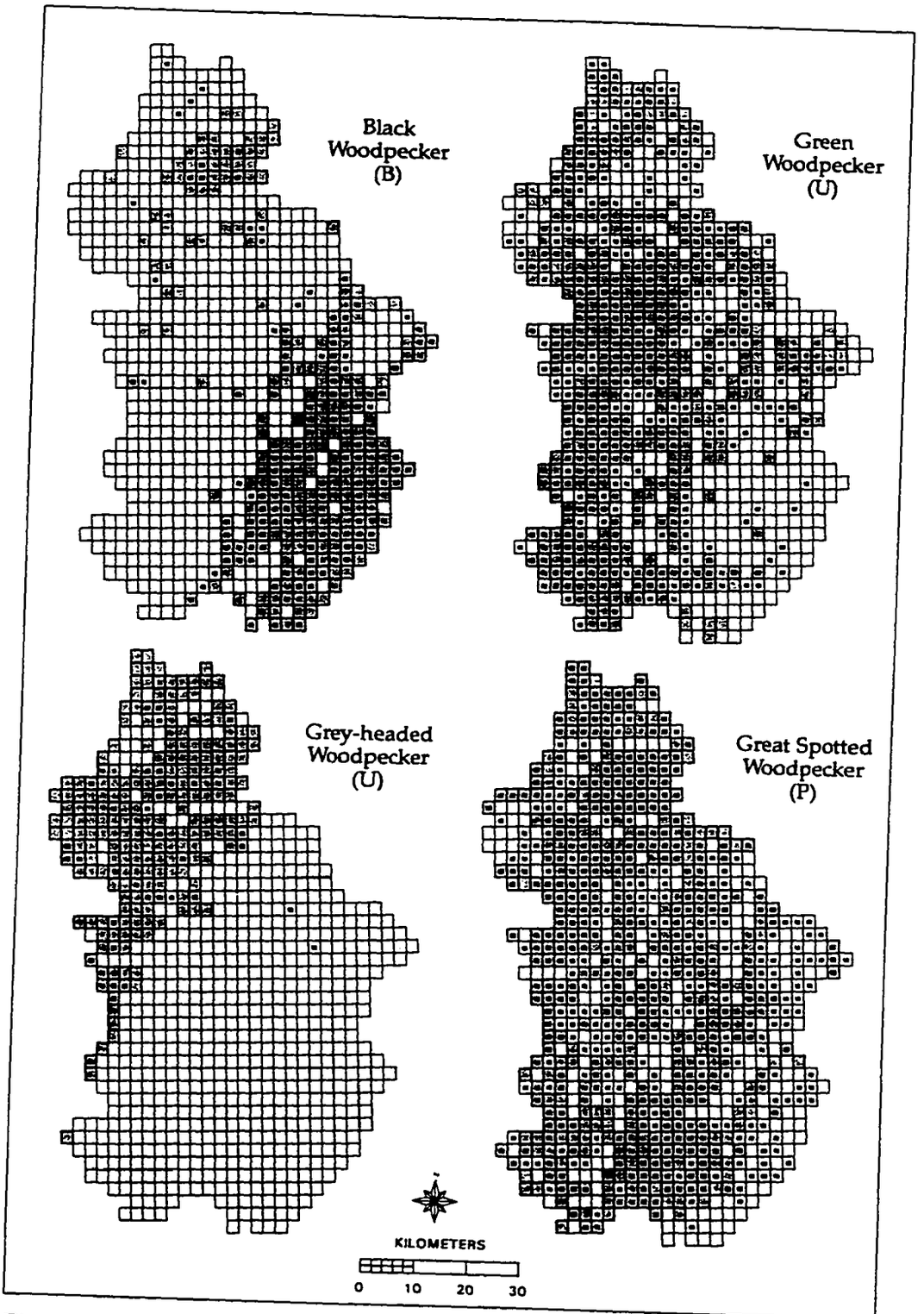


Figure 3. Predicted distribution of seven woodpecker species in the Jura, France, using logistic regression models. Shaded: predicted presence; dot: true presence. Only the best model (highest Kappa value) for each species is shown. B = biological, P = most parsimonious, U = universal.

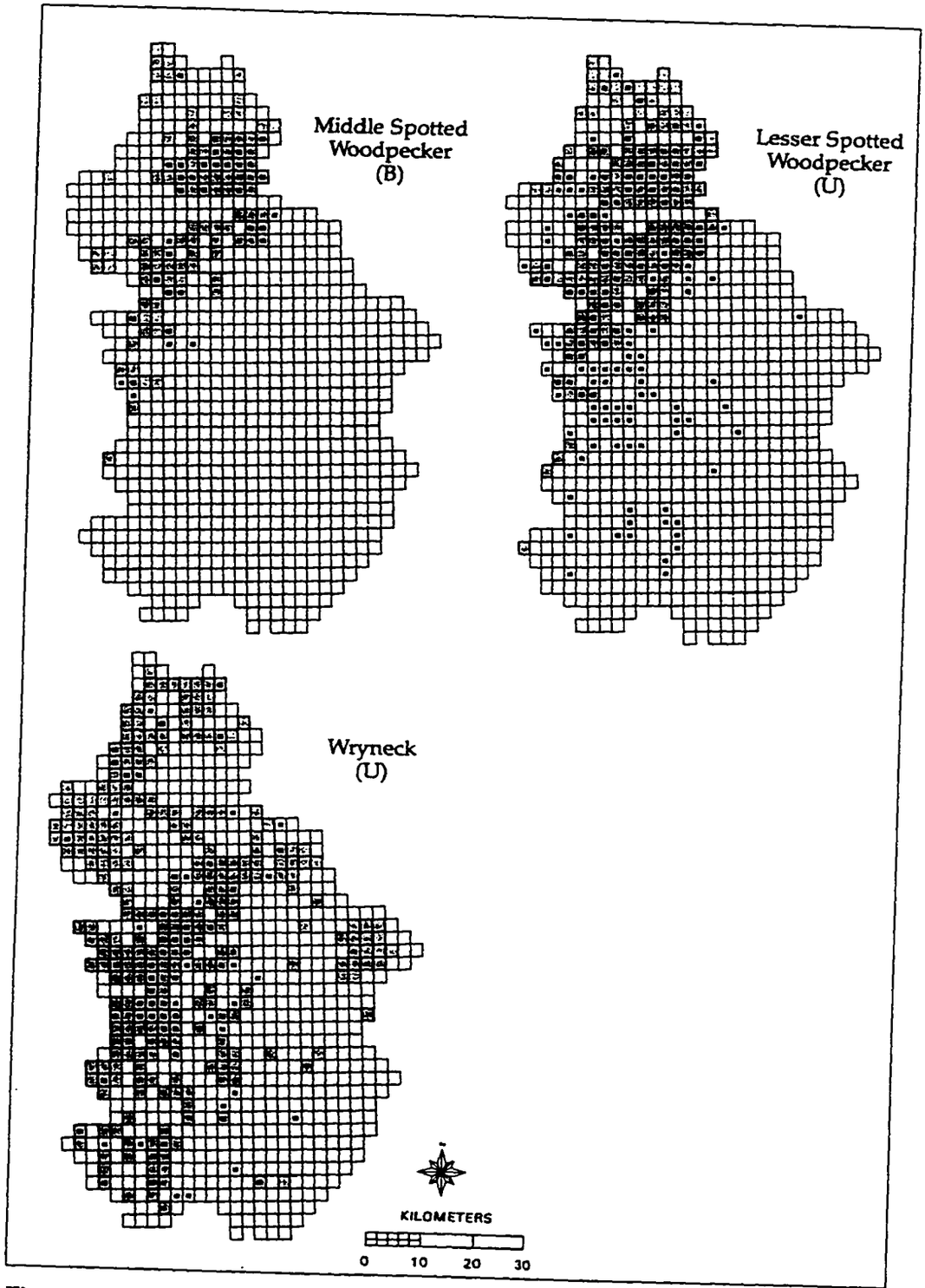


Figure 3 (continued). Predicted distribution of seven woodpecker species in the Jura, France, using logistic regression models. Shaded: predicted presence; dot: true presence. Only the best model (highest Kappa value) for each species is shown. B = biological, P = most parsimonious, U = universal.

DISCUSSION

Comparison of model types

Because I did not know what the performance of woodpecker-habitat relationship models created with broad-scale independent variables would be, I developed three types of models for each of the seven woodpecker species. This allowed an assessment of whether having more variables (while not overfitting the model) or having a more detailed habitat map affected the predictive ability of the models. Although the three model types performed differently, there was no general trend. For the black woodpecker, adding three forest classes (MtnDecid, Beech, ShelterCut) to a basic model (Fir, Spruce, Open) improved *K* by only 1.4%. For the middle spotted woodpecker, however, entering Meanelev and OtherCut in a model containing MatDecid only resulted in a 4.8% improvement. The opposite was observed for the lesser spotted woodpecker, for which a two-variables model was 7.5% better than a five-variables one. The better fit of the universal model for the green woodpecker may be due to the fact that the general forest class Decid included several forest types, such as coppice forests, that were not strongly correlated with the species' presence (based on univariate LR), but that could nonetheless have been used by the bird.

Generally speaking, *K* values for each species for the three models were in the same order of magnitude (Table 6). Thus, if my sole goal was to compare predicted presence with true ones, the benefit of including more forest classes or more variables would be slight, although more important for some species than others. The choice of an approach depends, in fact, on the desired use of the model. If models are developed to assess how sensitive the species' distributions are to changes in the variables' values, then all the variables that are thought to strongly influence the species' occurrence if they

were to be modified (e.g., change in the cutting regime), should be included in the model, regardless of their statistical significance (Fielding and Haworth 1995). This could be an argument for preferring the biological model of the black and middle spotted woodpeckers, for example. For both species, forest regeneration variables entered the biological models, where they negatively affected the species' probability of presence. Nevertheless, the relatively good results obtained with the universal models seem to indicate that, at a coarse scale of analysis, models can be built even if the forest cover map only differentiates among broad categories such as Decid, Mixed, Planted, and Open.

Comparing P_o and K values obtained for the woodpecker models shows that models can have a lower P_o, but provide a better improvement over chance classification, than others. This is the case for the green and grey-headed woodpeckers: K values are more indicative of the models' performances as suggested by the prediction maps. Based on the high percent correct classification rates obtained for the grey-headed woodpecker ($P_o > 75\%$; Table 6), a better correspondence between true and predicted presences and absences would be expected (Fig. 3g). The lower K values (below 18%; Table 6) obtained for this species are a much better indication of the poor fit of the models. The ten-fold validation procedure provided an additional criterion of the models' performances; poor stability (high CV) was obtained for the variables entering the wryneck' models, compared to other species.

Comparison of model performances across species

Although no one type of model performed systematically better than the other two, prediction accuracies varied considerably among species, indicating that certain

woodpeckers were more easily modeled than others at this scale, regardless of model type. Gates et al. (1995) obtained similar results with logistic models developed for eight farmland bird species. For example, 14 models explained between 28.1% and 33.8% of the deviance for the skylark (*Alauda arvensis*), and 12 models explained between 70.2% and 76% of the deviance for the turtle dove (*Streptopelia turtur*).

Because the habitat preferences of European woodpeckers have been studied quite thoroughly, I assessed whether or not the broad-scale variables I used were biologically meaningful, and provided tentative explanations for the good - or poor - fit of the models.

The black woodpecker prefers large forests of various composition (oak-beech and beech forests in the plains, beech-fir, beech and spruce forests at higher elevation), but mostly old and open beech forests, where the large diameter trees it requires for nesting are most likely to be found (Jovéniaux 1993). Cuisin (1988) observed a similar preference for big trees with smooth bark (beech, pine) in France and reported that in eastern Germany, 98% of the nests were found in beech trees. In the Rhine valley, Spitznagel (1990) hypothesized that the expansion of the black woodpecker from mixed and coniferous forests in montane ecosystems to lowland and pure deciduous forests, which has occurred since the turn of the century, is probably due to the introduction of beech into floodplain forests.

At the 575-ha scale of analysis in the Jura, presence of breeding black woodpeckers was associated with higher-elevation, relatively large mixed and coniferous forests (beech, fir, spruce), but was negatively correlated with deciduous forests, patchy woodlots, open areas, and shelterwood cuts (Table 2). Logically, large beech trees are most likely to be found in beech forests; but a positive correlation with fir forests, a species rarely chosen for nesting, was also observed (Mann-Whitney U-

test, $P < 0.001$, Table 2). Two factors account for this pattern. First, fir forests as defined by the IFN (1980) are never pure but contain a certain proportion of deciduous trees, mostly beech (IFN 1980, pers. obs.). Hence, a grid cell classified as 100% fir forest may contain several large-diameter beech trees suitable for nesting. Second, I found a significant positive correlation between the occurrence of beech and fir forests within the same cell (Spearman's $r = 0.36$, $P \leq 0.0005$, Table 4). A positive correlation between presence of the bird and beech forests will imply such a correlation with fir forests.

The black woodpecker is considered to be more of a generalist, in terms of habitat selection, than other members of the family (Wesolowski and Tomialojc 1986, Tjernberg *et al.*1993, Angelstam and Mikusinski 1994). In the Jura however, at the scale and for the variables considered, it appeared to be a somewhat specialized species, and it is the strong correlation between the bird's presence and certain forest types (such as Beech, Fir, and Spruce) that resulted in high Kappa values for the black woodpecker's models. The difference between generalist and specialist is scale-dependent: had my study area been all of France, the recent expansion of the bird in lowland deciduous forests would have suggested a generalist. Species' ranges are always fluctuating, but at different spatial and temporal scales. Accurate distribution models are particularly difficult to develop for species whose range is rapidly expanding (or retracting). For example, last century, Ogerien (1863, in Jovéniaux 1993) described the black woodpecker as a mountain species found in fir forests of the Upper Jura. Its sporadic use of lowland deciduous forests may indicate an expansion similar to that taking place at the larger scale, and is more difficult to model. Some presences in the deciduous Chaux forest could nevertheless be predicted (Fig. 2 and 3a) because the cells were almost entirely forested, and Open entered the models negatively. On the other

hand, if a presence was recorded in an atlas grid cell that was positioned to include mostly open habitat (for example, at the edge of the Chaux forest), the model did not predict a presence for that cell. Such a grid positioning problem is inherent to working with atlas data (Fielding and Haworth 1995).

The middle spotted woodpecker is typically a lowland bird (Cramp 1985), even in Switzerland where it does not normally breed above 700 m; in Canton Zurich, it appears restricted to the lowest woods (below 400 m; Glutz von Blotzheim and Bauer 1980). The univariate analysis I conducted revealed a strong association with low elevation and mature deciduous forests, and negative correlation with open areas and forest cover type richness. The middle spotted woodpecker was the second best predicted species, most certainly because at the scale of the study it appeared as a highly specialized bird strongly dependent on mature oak forests; the variable MatDecid alone correctly predicted 85.71% of presences and 87.58% of absences. This preference for large forests of mature oak trees is well known (Short 1982, Wesolowski and Tomialojc 1986, Spitznagel 1990, Jovéniaux 1993, Angelstam and Mikusinski 1994).

The green woodpecker is a ground species that feeds predominantly on ants and usually nests in parks, orchards, gardens, wooded pastures, and along edges (Short 1982, Cramp 1985, Clergeau and Chefson 1988, Seitz 1988, Spitznagel 1990, Angelstam and Mikusinski 1994). This species avoids dense, unbroken forests (Jovéniaux 1993, Angelstam and Mikusinski 1994), especially if they are coniferous (Cramp 1985), and seems to prefer low elevation, deciduous stands (Glue and Boswell 1994). The 575-ha scale of analysis confirmed the importance of edges, small forest patches, and open areas to the species, as well as its avoidance of dense coniferous forests. Its preference for low elevation was also identified, and the high upper limit (1236 m; Table 2) is consistent with observations in neighboring Switzerland (2100 m;

Glutz von Blotzheim and Bauer 1980). Good prediction rates (percent improvement over chance) were obtained, even though many forest patches that may contain suitable nesting habitat were not mapped at the 1:200,000 scale. This could be because what the IFN classified as Open is, in fact, rarely devoid of trees; instead, these are present in the form of hedgerows or small clumps mixed in the agricultural landscape (pers. obs.). Plotting the logit of the models against percent non forest resulted in a skewed Gaussian curve, i.e., probability of green woodpecker presence increased with increasing amount of non forest until non forest occupied about 90% of a cell. A linear, and not a Gaussian curve, would be expected if the non forest class always contained a few trees. Many cells containing 90% or more of open habitat were found in the more intensive agricultural areas, where hedgerows and groves have been removed; these cells should lack the forest elements required by the bird, and point to the intensification of agricultural practices as a threat to the species (Jovéniaux 1993).

The lower Kappa values obtained for the four remaining species can be partially explained by several factors. The great spotted woodpecker has been described as the most ubiquitous European woodpecker, both in terms of feeding requirements and habitat selection (Clergeau and Chefson 1988, Török 1990, Tjernberg *et al.* 1993, Angelstam and Mikusinski 1994), occupying forests of any structure and size as long as a few old trees are present (Cramp 1985, Jovéniaux 1993). My broad scale analysis in the Jura revealed an apparent avoidance of open areas, a result also observed in England (Hinsley *et al.* 1995). The great spotted woodpecker was the most common of the seven species (more presences than absences) and it occurred in every mapped forest class. This lack of strong habitat preferences, at the 575-ha scale, made its distribution difficult to predict. Although the models still performed better than chance, I suspect that if they were built using density of observation rather than just presence-absence data, they

might be more sensitive to small differences in habitat preferences, and thereby more accurate. For example, Jovéniaux (1993) found the species' densities to be much higher in mature deciduous forests (one to three pairs per ten ha in the deciduous Chaux forest, against one pair per 100 ha in the coniferous Massacre forest).

For the grey-headed and lesser spotted woodpeckers, and for the Wryneck, poor modeling results were related to scale. I found a correlation between the grey-headed woodpecker's presence and mature deciduous and coppice forests, as suggested by Angelstam and Mikusinski (1994). The variables used in the species' models (Meanelev, MatDecid, Coppice; Table 5) were biologically sound, but they predicted a much broader distribution than that of the original atlas map (Fig. 3g). This could be because at the level of individual forest patches, this woodpecker prefers more humid areas, such as beech, poplar, alder, or willow patches along streams and rivers (Cramp 1985, Jovéniaux 1993, Winkler *et al.* 1995). A GIS layer of hydrography may have improved the fit of the models by lowering the probability of presence in cells where no riparian features were present.

The lesser spotted woodpecker was associated with lower elevation, mature deciduous and coppice forests. Clergeau and Chefson (1988) described it as an open habitat species that prefers deciduous forests, and in Finland it is significantly more present in nemoral deciduous forest than in mixed woods (Olsson *et al.* 1992, Wiktander *et al.* 1992). Because it preferentially excavates highly decayed wood, it is often found in riparian forests (Cramp 1985, Jovéniaux 1993), where snags are often more common because of less intensive harvesting of the stands (Spitznagel 1990, Olsson *et al.* 1992). At the 575-ha scale, though, presences ranked significantly lower than absences for the riparian and heterogeneous forest class (Mann-Whitney U-test, $P < 0.001$, Table 2). A probable explanation is that riparian habitat makes up only a small fraction of the area

covered by this forest class, most of it being wooded pastures. Many small riparian patches have probably not been mapped at the 1:200,000 scale of the IFN map, but when true riparian habitat can be identified in an atlas cell (i.e., when the class patches are located along rivers), the species was often recorded in that cell. I suspect that subdividing the riparian and heterogeneous forest class into its riparian and wooded pasture components may have lead to a better map of preferred habitat and in turn lead to a better model fit. As is, the models incorporate only such broad forest types as MatDecid or Coppice, causing type I (or omission) errors (presences not predicted) where small forest patches with decayed wood were not mapped, and type II (or commission) errors (absences predicted as presences) where one of these two types occurred, even though it may not contain trees suitable for nest excavation (Fig. 3e).

Finally, the Wryneck is known to prefer sunny, open sites over closed coniferous stands (Short 1982, Cramp 1985, Jovéniaux 1993, Winkler *et al.* 1995). The species is considered a low elevation species, although it may occasionally nest at high altitude (2100 m in a Trentino valley, Italy; Pedrini 1984). The univariate test identified preferences for patchy forests and open areas, and the avoidance of fir and spruce forests. The Wryneck is currently declining in the Jura, partially because of small-scale habitat changes such as the disappearance of small patches of dry grasslands, of orchards, and even the felling of individual trees (Jovéniaux 1993). Clearly, a 1:200,000 scale forest map cannot capture such fine-scale habitat characteristics, explaining the poor fit of the model.

Using atlas data to obtain bird-habitat relationship information

Even though the correlations I observed between woodpecker presence and habitat variables are in general accordance with what is known of the species' habitat

preferences, working with broad-scale distribution data presents limitations. A number of studies demonstrate that patterns of bird-habitat relationships vary with the scale of investigation (e.g. Wiens *et al.* 1987, Orians and Wittenberger 1991, Brandt *et al.* 1995). The nature of my data did not allow analysis at a smaller scale (i.e., nest location data were not available), but contrasting the existing knowledge on small-scale habitat relationships with the results of the larger-scale approach could provide insights into scale-dependent patterns of habitat associations for woodpecker species. A good example is provided by the distribution of the black woodpecker relative to patches of shelterwood cuts and other types of cuts. Based on univariate results, the species appeared to avoid the former, but not the latter (Table 2). This is surprising, considering that its breeding requirements at the local scale (big trees) are not compatible with a young, even-aged fir or spruce plantation, or an open clearcut (although the bird does feed in open areas). A look at the spatial distribution and characteristics of the two types of regeneration cut helps to clarify the apparent contradiction. Shelterwood cuts are generally located at low to medium elevation, in forest types poorly colonized by the black woodpecker (mature deciduous or coppice), and are usually large tracts (156 patches, mean size = 84.5 ha, maximum = 2480 ha in the Moidons forest). Other cutting types are found at all elevations; at high elevation, they appear as small inclusions (527 patches, mean size = 32.6 ha, maximum 640 ha in the Poligny forest) within older mixed forests that were positively correlated with the bird's presence (beech, fir and spruce). Because forest composition was extracted from 575-ha grid cells, small patches of forest types negatively associated with the bird's breeding presence at the local scale have little effect on the presence of the bird within a cell. On the other hand, small clearcuts are probably beneficial to the species because they provide foraging habitat (Clergeau and Chefson 1988, Rolstad *et al.* 1998, pers. obs.).

In Sweden, forest fragmentation by small clearcuts (1.4 to 4.8 ha) did not appear to affect the species (Tjernberg *et al.* 1993). Shelterwood cuts, on the other hand, can cover a large portion of a cell's area. Although this cover type may be used for foraging, it probably does not offer suitable nesting characteristics. This example illustrates that knowledge of smaller scale information is necessary to interpret or assess the validity or broad-scale patterns. Parker (1996) reached a similar conclusion when interpreting the results of bird-habitat models developed in one-eighth degree squares in Swaziland: the significant, positive association between the bald ibis *Geronticus calvus* and exotic timber plantations only indicated that the species occurred in a habitat that was suitable for the establishment of plantations -- not that the bald ibis benefitted from them.

Atlas data and landscape variables

The habitat variables I used were simple and 'typical' of any large-scale wildlife-habitat relationship study, especially when grid cells are used as units. The spatial patterning of habitat types across landscapes, however, is thought to exert a strong influence on the distribution of the vertebrate populations inhabiting those landscapes (Wiens 1989). The first studies on habitat selection conducted at the landscape scale originated from the extension of the theory of island biogeography (MacArthur and Wilson 1967) to 'terrestrial islands', i.e., forest fragments isolated in a 'sea' of open lands (Harris 1984). The effect of patch size on composition and richness of avian communities has been the focus of several studies (e.g., Forman *et al.* 1976, Whitcomb *et al.* 1981, Lynch and Whigham 1984). Recent developments in GIS and spatial statistics make it possible to investigate the effect of spatial patterning of a landscape on habitat selection (e.g., Hansen and Urban 1992, Gustafson et al 1994,

Lescourret and Genard 1994, McGarigal and McComb 1995). In these studies, a variety of landscape indices were computed to quantify various aspects of landscape configuration (e.g. area, shape, nearest-neighbor, diversity indices) and correlations with species distribution, richness or abundance were measured. Woodpeckers might be area-sensitive (Cramp 1985, Petterson 1985, Haila *et al.* 1987, Wiktander *et al.* 1992, Hinsley *et al.* 1995), although black woodpeckers can do well if forest fragments with a certain proportion of large-diameter trees occupies a sufficient proportion of the landscape (Cramp 1985, Tjernberg *et al.* 1993). The relationship between the distribution of woodpecker species and spatial characteristics of the landscape other than patch area and isolation has not been studied.

Unfortunately, bird atlas data are poorly suited to assess the influence of spatial characteristics of the landscape on the distribution of species. Presence-absence data are recorded on a cell-by-cell basis, making it impossible to obtain basic spatial information such as patch size: each cell usually overlaps many patches, so it is not possible to know which one to consider. Computing mean patch size within cells could lead to confusing results, because a cell with several, medium-sized patches could give the same outcome as one with one large patch and many small ones. Similarly, nearest-neighbor and patch type adjacencies cannot be obtained. A large number of other landscape indices can be obtained through FRAGSTATS (McGarigal and Marks 1994), a spatial analysis program. I tried to run it on test cells, but the coarse resolution of the forest map compared to the grid cell size rendered the results meaningless, i.e., there were not enough forest polygons per cell to have a reasonable sample size. Edge density and forest class richness were the only two landscape variables that could be easily calculated from the data. Number of patches was strongly correlated with Richness (Spearman's $r = 0.75$, $P \leq 0.0005$, Table 4) and Edgeden (Spearman's $r =$

0.88, $P \leq 0.0005$, Table 4) and thus was not included in the analysis. Percent of forested landscape was indirectly obtained by percent Open, which turned out to be negatively associated with the presence of forest woodpeckers (black, great and middle spotted woodpeckers) and positively associated with that of species known to occur in more open landscapes (green woodpecker and Wryneck). I could not verify, however, whether middle spotted woodpeckers preferred patches of 30 ha or more, and avoided those below 5 ha (Cramp 1985). The negative association between the species' presence and edge density (Table 2) is the only indication of the species' possible avoidance of small patches.

Certain variables lose their meaning when computed over a large area (i.e., in a 575-ha cell). In my study, this was the case for Slope. In addition to being strongly correlated with Meanelev, this variable was not relevant to the species' ecology at the landscape scale of analysis. At a finer scale, for example if individual nest sites were mapped, it may reveal more interesting trends. Slope probably has little influence on the species *per se*, but certainly affects the way forests are managed. In the Upper Jura, I noticed several instances when black woodpeckers had excavated their nests in beech trees growing on very steep slopes (e.g., greater than 45 degrees). Slope could be a predictor of nest site location by identifying areas where large beech trees are more likely to be left uncut.

Correlations and models derived from atlas data as a management tool for woodpeckers in the Jura

Despite the limitations of atlas data outlined above, important information on habitat relationships was extracted for woodpecker species from the Jura breeding bird atlas. Quantitative data obtained for the entire *département* can be compared with

anecdotal information. For example, the conversion of deciduous forests to conifer ones has been suggested as a factor potentially contributing to the decline of the middle spotted woodpecker (Pettersson 1985, Clergeau and Chefson 1988). Even though the univariate analysis did not reveal a statistically significant difference between presences and absences for variables such as ShelterCut and OtherCut (Table 2), the maximum percentages of these two classes in cells where the bird was breeding were less than they were for any of the other six species (12% and 13%, respectively; Table 2). Because the biological model reflects the importance of mature deciduous forests to the species, and includes the variable OtherCut, it could be used to simulate the consequences of additional conversion on the species' breeding distribution. Spatial simulation modeling may be one of the most powerful applications of GIS in future land and resource management (Parker 1988); it would be simple to modify the existing forest map, then run the models, and compare the resulting distribution map to that obtained with the original forest map. Because the models were not validated with an independent data set, it would be difficult to assess how much confidence should be put in the output of such a simulation (Morrison *et al.* 1992). Simulation modeling could nevertheless be used as an exploratory tool, as a 'warning signal' of what may happen under various management procedures, and lead forest managers to think more critically about the consequences of broad-scale modifications of the landscape. The conversion of much of the First Plateau deciduous forests to uniform conifer stands after World War II is an example of such an alteration of the semi-natural landscape.

Another possible use of models developed from atlas data is to improve the atlas itself. Because the output of logistic regression is not categorical (i.e., presence and absence), but probabilistic, the models can be used to highlight cells with a higher probability of species' presence. This was done in Lesotho, where the difficulty to

access the most rugged parts of the country and an uneven repartition of the number of observers may have biased species distribution estimates (Osborne and Tigar 1992). During the creation of the Jura atlas, much time and effort was devoted to insure as thorough a coverage of the *département* as possible. Nevertheless, censuses may have underestimated the distribution of such secretive and rare species as the grey-headed and middle spotted woodpeckers (Jovéniaux pers. commun.). Cells with high probabilities of presence but where no bird was detected (Fig. 3) would be very useful to target future censuses. In addition, because population sizes of species fluctuate, it is necessary to identify and protect areas of suitable habitat, even if they are not occupied at the moment; this is especially important for species that have declined but may be in a recolonizing phase (Anthony *et al.* 1982). Identifying which cells have a high probability of presence would help biologists prioritize forests where more traditional, finer-scale, time-consuming habitat investigations should be conducted.

Overall, my study showed that analyzing woodpecker distribution data from a breeding bird atlas in conjunction with habitat variables in a GIS provides a relatively simple way to build useful habitat relationships models. With more species becoming threatened by alteration of their environment under human pressure, and little time or funding available for detailed, fine-scale studies, it will become crucial to make the best use of existing information. Models derived from atlas data could be valuable tools to contribute to the long-term protection of wildlife species -- as long as one is aware of the limitation of working at a broad scale and with grid-based data.

REFERENCES

Angelstam, P. 1990. Factors determining the composition and persistence of local woodpecker assemblages in taiga forest in Sweden - a case for landscape ecological studies. Pages 147-164 *in* A. Carlson and G. Aulén, eds. Conservation and Management of Woodpecker Populations. Report 17, Swedish Univ. of Agric. Sci., Uppsala.

Angelstam, P, and G. Mikusinski. 1994. Woodpecker assemblages in natural and managed boreal and hemiboreal forest - a review. Ann. Zool. Fennici **31**, 157-172.

Anthony, R. G., R. L. Knight, G. T. Allen, B. R. McClelland, and J. I. Hodges. 1982. Habitat use by nesting and roosting bald eagles in the Pacific Northwest. Trans. N. Amer. Wildl. and Natur. Resour. Conf. **47**, 332-342.

Bobbink, R., and J. H. Willems. 1993. Restoration management of abandoned chalk grassland in the Netherlands. Biodiversity and Conserv. **2**, 616-626.

Böhning-Gease, K., and H. G. Bauer. 1996. Changes in species abundance, distribution, and diversity in a Central European bird community. Conserv. Biol. **10**, 175-187.

Brandt, C. A., J. K. Parrish, and C. N. Hodges. 1995. Predictive approaches to habitat quantification: dark-rumped petrels on Haleakala, Maui. Auk **112**, 571-579.

Capen, D. E., J. W. Fenwick, D. B. Inkley, and A. C. Boynton. 1986. Multivariate models of songbird habitat in New England forests. Pages 171-175 *in* J. Verner, M. L. Morrison, and C. J. Ralph, eds. Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates. Univ. of Wisconsin Press, Madison.

Clergeau, P., and P. Chefson. 1988. Les Pics. Série "Comment vivent-ils" Vol. 20. Atlas Visuels Payot Lausanne. Payot Lausanne, Lausanne.

Cramp, S. 1985. The Birds of Western Palaearctic, Vol. IV. Oxford Univ. Press, Oxford.

Cuisin, M. 1980. Nouvelles données sur la répartition du pic noir (*Dryocopus martius* (L.)) en France et comparaison avec la situation dans d'autres pays. L'oiseau et R.F.O. **50**, 23-32.

Cuisin, M. 1988. Le pic noir (*Dryocopus martius* (L.)) dans les biocénoses forestières. L'oiseau et R.F.O. **58**, 173-276.

Ehrlich, P. R., D. S. Dobkin, D. Wheye, and S. L. Pimm. 1994. The Birdwatcher's Handbook - A Guide to the Natural History of the Birds of Britain and Europe. Oxford Univ. Press, Oxford.

ESRI. 1994. PC Arc/Info 3.4.2. Redlands, CA, U.S.A.

Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conserv. Biol. **9**, 1466-1481.

Forman, R. T. T., A. E. Galli, and C. F. Leck. 1976. Forest size and avian diversity in New Jersey woodlots with some land use implications. Oecologia **26**, 1-8.

Gates, S., D. W. Gibbons, P. C. Lack, and R. J. Fuller. 1993. Declining farmland bird species: modelling geographical patterns of abundance in Britain. Pages 153-177 *in* P. J. Edwards, R. M. May, and N. R. Webb, eds. Large-Scale Ecology and Conservation Biology. Blackwell Sci. Publ., London.

Gauch, H. G., and G. B. Chase. 1974. Fitting the Gaussian curve to ecological data. Ecology **55**, 1382-1390.

Glayre, D., and D. Magnenat. 1984. Oiseaux Nicheurs de la Haute Vallée de l'Orbe. Nos Oiseaux 398. Fascicule spécial du volume 37.

Glue, D. E., and T. Boswell. 1994. Comparative nesting ecology of the three British breeding woodpeckers. British Birds **87**, 253-269.

Glutz von Blotzheim, U. N., and K. M. Bauer. 1980. Handbuch der Vögel Mitteleuropas. Band 9. Akademische Verlagsgesellschaft, Wiesbaden.

Green, R. E., P. E. Osborne, and E. J. Sears. 1994. The distribution of passerine birds in hedgerows during the breeding season in relation to characteristics of the hedgerow and adjacent farmland. J. Appl. Ecol. **31**, 677-692.

Gustafson, E. J., G. R. Parker, and S. E. Backs. 1994. Evaluating the spatial pattern of wildlife habitat: a case study. Am. Midl. Nat. **131**, 24-33.

Haila, Y., I. K. Hanski, and S. Raivio. 1987. Breeding bird distribution in fragmented coniferous taiga in southern Finland. Ornis Fennica **64**, 90-106.

Hansen, A. J., and D. L. Urban. 1992. Avian response to landscape pattern: the role of species' life history. Landscape Ecol. **7**, 163-180.

Harris, L. D. 1984. The Fragmented Forest. Univ. of Chicago Press, Chicago.

Hinsley, S. A., P. E. Bellamy, I. Newton, and T. H. Sparks. 1995. Habitat and landscape factors influencing the presence of individual breeding bird species in woodland fragments. J. of Avian Biol. **26**, 94-104.

Hosmer, D. W., and S. Lemeshow. 1989. Applied Logistic Regression. John Wiley and Sons, New York.

IFN 1980. Département du Jura. Résultats du Deuxième Inventaire Forestier. Tome 1. Ministère de l'Agriculture, Direction des Forêts.

IFN 1993. Carte Forestière du département du Jura. Ministère de l'Agriculture, Direction des Forêts.

James, F. C., and C. E. McCulloch. 1990. Multivariate analysis: panacea or Pandora's box? Annu. Rev. Ecol. Sys. **21**, 129-166.

Johnson, L. B. 1990. Analyzing spatial and temporal phenomena using geographic information systems: a review of ecological applications. Landscape Ecol. **4**, 31-43.

Johnsson, K, S. G. Nilsson, and M. Tjernberg. 1990. The black woodpecker - a key species in European forests. Pages 99-102 *in* A. Carlson and G. Aulén, eds. Conservation and Management of Woodpecker Populations. Report 17, Swedish Univ. of Agric. Sci., Uppsala.

Johnsson, K., S. G. Nilsson, and M. Tjernberg. 1993. Characteristics and utilization of old black woodpecker *Dryocopus martius* holes by hole-nesting species. Ibis **135**, 410-416.

Jovéniaux, A. 1993. Atlas des Oiseaux Nicheurs du Jura. G.O.J. ed., Lons-le-Saunier.

Lescourret, F., and M. Genard. 1994. Habitat, landscape and bird composition in mountain forest fragments. J. Environ. Manage. **40**, 317-328.

Livingston, S. A., C. S. Todd, W. B. Krohn, and R. B. Owen, Jr. 1990. Habitat models for nesting bald eagles in Maine. J. Wildl. Manage. **54**, 644-653.

Loftsgaarden, D. O., and P. L. Andrews. 1992. Constructing and testing logistic regression models for binary data: applications to the National Fire Danger Rating System. USDA Forest Service, Intermoutain Research Station General Technical Report INT-286.

Lynch, J. F., and D. F. Whigham. 1984. Effects of forest fragmentation on breeding

bird communities in Maryland, U.S.A. Biol. Conserv. **28**, 287-324.
MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton Univ. Press, Princeton.

McCullagh, P., and J. A. Nelder. 1983. Generalized Linear Models. Chapman & Hall, London.
McGarigal, K, and B. J. Marks. 1994. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. User's manual version 2.0. Oregon State Univ., Corvallis.
McGarigal, K, and McComb, W. C. 1995. Relationships between landscape structure and breeding birds in the Oregon Coast Range. Ecol. Monographs **65**, 235-260.
Mikusinski, G., and P. Angelstam. 1998. Economic geography, forest distribution, and woodpecker diversity in Central Europe. Conserv. Biol. **12**, 200-208.
Morrison, D. G. 1969. On the interpretation of discriminant analysis. J. Marketing Res. **4**, 156-163.
Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1992. Wildlife-Habitat Relationships. Univ. of Wisconsin Press, Madison.
Moss, D., P. N. Taylor, and N. Easterbee. 1979. The effects on song-bird populations of upland afforestation with spruce. Forestry **52**, 129-147.
Olsson, O., I. N. Nilsson, S. G. Nilsson, B. Pettersson, A. Stagen, and U. Wiktander. 1992. Habitat preferences of the lesser spotted woodpecker *Dendrocopos minor*. Ornis fennica **69**, 119-125.
ONF (1993). Instruction "biodiversité". 93 T 23 du 15 novembre 1993.
Orians, G. H., and J. F. Wittenberger. 1991. Spatial and temporal scales in habitat selection. Am. Nat. **137** (suppl.), S29-S49.
Osborne, P. E., and B. J. Tigar. 1992. Interpreting bird atlas data using logistic models: an example from Lesotho, Southern Africa. J. Appl. Ecol. **29**, 55- 62.
Parker, H. D. 1988. The unique qualities of a geographic information system: a commentary. Photogrammetric Engineering and Remote Sensing **54**, 1547-1549.
Parker, V. 1996. Modeling the distribution of bird species in Swaziland in relation to environmental variables. Ostrich **67**, 105-110.
Pedrini, P. 1984. Nesting of the gray-headed woodpecker, *Picus canus*, at high altitude. Rivista Italiana di Ornithologia **54**, 266-267.
Pereira, J. M. C., and R. M. Itami. 1991. GIS-based habitat modeling using logistic regression: a study of the Mt. Graham red squirrel. Photogrammetric Engineering and Remote Sensing **57**, 1475-1486.
Pettersson, B. 1985. Extinction of an isolated population of the middle spotted woodpecker *Dendrocopos medius* (L.) in Sweden and its relation to general theories on extinction. Biol. Conserv. **32**, 335-353.
Robertson, A., R. E. Simmons, A. M. Jarvis, and C. J. Brown. 1994. Can bird atlas data be used to estimate population size? A case study using Namibian endemics. Biol. Conserv. **71**, 87-95.
Rolstad, J., P. Majewski, and E. Rolstad. 1998. Black woodpecker use of habitat and feeding substrates in a managed Scandinavian forest. J. Wildl. Manage. **62**:11-23.
Schifferli, A., P. Géroudet, and R. Winkler. 1980. Atlas des Oiseaux nicheurs de Suisse. Station ornithologique Suisse, Sempach.

Seitz, B. J. 1988. The relationship between vegetation complexes and bird communities in cultivated land: studies in southwestern German hilly land. Phytocoenologia **16**, 315-390.

Short, L. L. 1982. Woodpeckers of the World. Delaware Museum of Natural History Monograph Series No. 4. Delaware Museum of Natural History, Greenville.

Sokal, R. R., and N. L. Oden. 1978. Spatial autocorrelation in biology. 1. Methodology. Biol. J. of the Linnean Soc. **10**, 199-228.

Spitznagel, A. 1990. The influence of forest management on woodpecker density and habitat use in floodplain forests of the Upper Rhine Valley. Pages 177-146 *in* A. Carlson and G. Aulén, eds. Conservation and Management of Woodpecker Populations. Report 17, Swedish Univ. of Agric. Sci., Uppsala.

SPSS, Inc. 1990. SPSS version 4.0.4. for the Macintosh. Chicago.

Stowe, T. J., A. V. Newton, R. E. Green, and E. Mayes. 1993. The decline of the corncrake *Crex crex* in Britain and Ireland in relation to habitat. J. Appl. Ecol. **30**, 53-62.

Taianen, J. 1985. Monitoring bird populations in Finland. Ornis Fennica **62**, 80-89.

Thomas, C. D., and J. C. G. Abery. 1995. Estimating rates of butterfly decline from distribution maps: the effect of scale. Biol. Conserv. **73**, 59-65.

Titus, K., J. A. Mosher, and B. K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: ecological applications. Am. Midl. Nat. **111**, 1-7.

Tjernberg, M., K. Johnsson, and S. G. Nilsson. 1993. Density variation and breeding success of the black woodpecker *Dryocopus martius* in relation to forest fragmentation. Ornis Fennica **70**, 155-162.

Török, J. 1990. Resource partitioning among three woodpecker species Dendrocopos spp. during the breeding season. Holarctic Ecol. **13**, 257-264.

Underhill, L. G., R. P. Prys-Jones, J. A. Harrison, and P. Martinez. 1992. Seasonal patterns of occurence of Palaearctic migrants in southern Africa using atlas data. Ibis **134** (Suppl.1), 99-108.

Wesolowski, T., and L. Tomialojc. 1986. The breeding ecology of woodpeckers in a temperate primaeval forest - preliminary data. Acta Ornithologica **22**, 1-21.

Whitcomb, R. F., C. S. Robbins, J. F. Lynch, B. L. Whitcomb, M. K. Klimjievicz, and D. Bystrak. 1981. Effects of forest fragmentation on avifauna of the eastern deciduous forest. Pages 123-205 *in* R. L. Burgess and D. M. Sharpe, eds. Forest Island Dynamics in Man-dominated Landscapes. Springer-Verlag, New York.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecol. **3**, 385-397.

Wiens, J. A., J. T. Rottenberry, and B. Van Horne. 1987. Habitat occupancy patterns of North American shrubsteppe birds: the effects of spatial scale. Oikos **48**, 132-147.

Wiktander, U., I. N. Nilsson, S. G. Nilsson, O. Olsson, B. Pettersson, and A. Stagen. 1992. Occurence of the lesser spotted woodpecker *Dendrocopos minor* in relation to area of deciduous forest. Ornis Fennica **69**, 113-118.

Winkler, H., D. A. Christie, and D. Nurney. 1995. Woodpeckers. An Identification Guide of the Woodpeckers of the World. Houghton Mifflin Company, Boston.

CHAPTER 2

TESTING THE UNIVERSALITY OF WOODPECKER-HABITAT RELATIONSHIP MODELS

Key words: habitat model, model universality, Switzerland, validation, woodpeckers.

ABSTRACT

To test the ability of habitat models developed for seven woodpecker species in the Jura, France, to predict woodpecker species distribution elsewhere, I applied the models to two sites in Switzerland for which woodpecker distribution maps were available. Because a forest cover map similar to that used for model development was not available for the test sites, I extracted land cover characteristics from a classified Landsat TM image. I increased the minimum mapping unit (MMU) of the Landsat TM image so that it was closer to that of the original forest cover map. Model performances, as measured by Cohen's Kappa and visual comparison of predicted vs true presences, were low, regardless of species, study site, and MMU of the land cover map. A comparison of the original forest cover map with the classified Landsat TM image was conducted for an area of overlap and revealed important differences between the two datasets, the most probable reason for the poor ability of the models to predict woodpecker species distribution in Switzerland. Additional factors that may have contributed to model failure are presented. I suggest that if models are developed with the goal of prediction, predictor variables should be dataset-independent and available outside the area used for model development.

INTRODUCTION

Wildlife-habitat relationship (WHR) models that relate the presence of wildlife species to characteristics of their environment have become a tool commonly used in wildlife management and conservation (Capen 1981, Verner *et al.* 1986). WHR models are developed to understand which factors affect a species' distribution, or to predict the distribution or abundance of species under conditions different from those used for model development - in a different area, or at a different time. With the advent of 'canned' statistical packages, statistical models, usually created using a mutlivariable approach, are a popular form of WHR models (Shugart 1981). Models, however, are developed faster than they are applied (Chalk 1986), and they often lack a proper validation test (Noon 1986). Validation should be an integral part of model development, because it indicates how much confidence can be placed in the outcome of the model (Morrison *et al.* 1992).

Marcot *et al.* (1983) provided a list of criteria useful for validating WHR models, from mathematical-based criteria to more subjective notions such as model appeal and credibility. Most validation techniques, however, emphasize accuracy, i.e., the match between a model's prediction and reality. Resampling procedures, which consist in using the same data for model development and validation, are one way to evaluate the classification accuracy of WHR models (Verbyla and Litvaitis 1989; see Chapter 1), but whenever possible, validation using an independent data set is recommended (Capen *et al.* 1986, Noon 1986, Morrison *et al.* 1992, Fielding and Bell 1997). One way to provide this independent set is to split the data into two sets and use one for model development, the other for validation (e.g. Chandler *et al.* 1995, Nadeau *et al.* 1995). Although this commonly used procedure is better than straight

resubstitution, where the same data are used for developing and testing the model, splitting data arbitrarily may not be the same as collecting new data (Chatfield 1995). For example, data partitioning does not fully address the question of the universality of a model, because the species data (presence/absence, population density, etc.) and habitat variables used in model development and model testing are likely to show little variability, being collected by the same person(s) using identical methods. It is conceivable, however, that the source of habitat variables used by the person applying the model may differ from that used in model development. This is especially likely if the original variables were extracted from a map, because an equivalent map may not be available elsewhere. In this case, one must wonder whether the model should be used at all, or whether it is general enough to perform reliably despite differences between data sets. The more different the data, the more likely it is that the model will perform poorly.

In Chapter 1, logistic regression models for seven woodpecker species in the Jura *département*, France, were built from simple, broad-scale habitat variables and atlas distribution maps of presences and absences. Three types of models were developed for each species: 'biological' models included statistically significant predictor variables, plus independent variables deemed important to the ecology of the species; 'parsimonious' models only included statistically significant independent variables; and 'universal' models were developed after simplifying the forest cover map into broad categories (such as deciduous or coniferous). In this chapter, I test the ability of these models to predict the distribution of woodpecker species in a new location, and with a data set different from that used for model development.

STUDY AREAS

I applied the three types of models to two test sites, both located in Switzerland: the upper Orbe Valley and Geneva Canton (Fig. 1). Landscape characteritics of the upper Orbe Valley are typical of high elevation valleys of the Jura mountain range: pastures occupy the valley bottom, whereas the slopes are covered with large, unbroken coniferous forests dominated by spruce (*Picea abies*) (Fig. 2). Geneva Canton, on the other hand, is located in the Geneva plain, where small, essentially deciduous woodlots are scattered in an agricultural landscape dominated by fields and crops (Fig. 2).

METHODS

Digital database

I obtained distribution maps of woodpecker presences and absences from the Breeding Bird Atlas of the Upper Orbe Valley (Glayre and Magnenat 1984) and from the Breeding Bird Atlas of Geneva Canton (Géroudet *et al.* 1983). Both atlases used 1-km^2 (i.e., 100 ha) grid cells, with 273 cells censused in the Orbe Valley, and 306 cells censused in Geneva Canton (Fig. 2). I used the Geographic Information System (GIS) software Arc/Info 7.0.3. (ESRI 1995) on a Unix workstation to recreate the atlas grids. I coded as "presence" all the cells in which woodpecker presence was labelled as Probable or Certain in the atlases (see Chapter 1 for justifications for this coding). Not all seven species of woodpecker found in the Jura were present in the two study areas. Six species were censused in Geneva Canton: the green woodpecker (*Picus viridis*), grey-headed woodpecker (*P. canus*), great spotted woodpecker (*Dendrocopos major*),

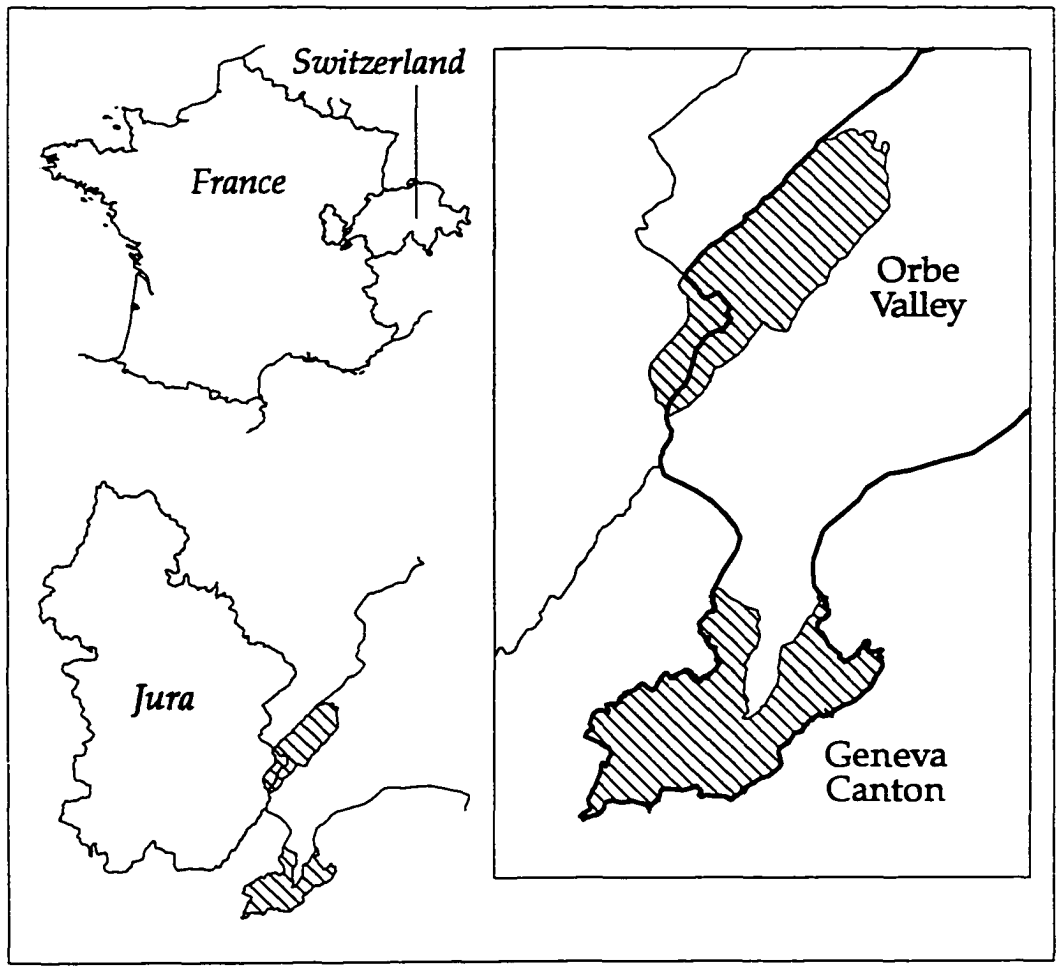


Figure 1. Location of the two study areas in relation to the Jura.

Figure 2. Distribution of cover types in the two study areas, and position of the two bird atlas grids. One grid atlas cell = 1 sq km.

middle spotted woodpecker (*D. Medius*), lesser spotted woodpecker (*D. minor*), and wryneck (*Jynx torquilla*), whereas only three were located in the Upper Orbe Valley: the black woodpecker (*Drycopus martius*) and the green and great spotted woodpeckers.

The models required three types of variables for each atlas cell: vegetation (% cover); edge density (m/ha); and mean elevation (m). In the absence of a land cover map similar to that used for model development (forest cover map from the Inventaire Forestier National - IFN; Chapter 1), land cover data were extracted from a 1992 classified Landsat TM image with a 25 m-pixel resolution (Vuillod 1994). The supervised classification was labelled by intensive ground-truthing of a topographically diverse 15 by 15 km^2 area and by comparison with existing, fine-scale land cover maps (Vuillod 1995), and resulted in 13 land cover classes (Table 1). Because the land cover classes used to develop the models in the Jura (Chapter 1) were different from those mapped in Switzerland from the TM image, it was necessary to combine types into more general, yet comparable classes (Table 1). The vegetative composition (%) of each atlas cell was obtained by overlaying the atlas grids with the classified image in the GIS.

Edge density (m/ha) was computed for each cell after grouping the forest cover types into one class, and the remaining classes into another. The resulting, simplified raster file was vectorized, and edge density was extracted for each cell following the procedure described in Chapter 1. Mean elevation (m) for each cell was extracted from a 50-m digital elevation model (DEM) purchased from the French Institut Géographique National. In the original model, elevation came from a 600-m DEM (Chapter 1).

Modeling procedures

Because there was no one-to-one correspondance between the IFN and the Landsat variables, I used my knowledge of the two study areas to modify the regression

Table 1. Landsat satellite image classification used to apply models developed for seven woodpecker species in the Jura, France, to two sites in Switzerland.

Original Landsat classes	Classes used for Biological and Parsimonious models	Classes used for Universal models
Xeric forests	N/A	N/A
Oak/hornbeam/beech forests	Decid	Decid
Beech forests	Beech	Mixed
Beech/fir/spruce forests	Mixed	
Coniferous plantations	Planted	Planted
Scrub/shrub	Open	Open
Pastures and fields		
Crops		
Low urban development		
Dense urban development		
Parkings, mines, quarries		
Bare rocks		
Water		

equations to fit the Landsat classes. For example, in the Orbe Valley coniferous forests are dominated by spruce, so when both fir and spruce entered the regression equation, I eliminated fir from the equation. When it was not clear which variable should be dropped, I tried running the model with each variable alternatively, and retained the model that worked best. For example, MatDecid and Coppice entered the biological model for the great spotted woodpecker, but the Landsat classification only had the general class Deciduous. In one model, the variable MatDecid was dropped; in the other, Coppice was dropped.

Once all models were run, I evaluated and compared their performances using the Kappa statistic (Titus *et al.* 1984), a suitable assessor of modeling success when the number of presences and absences do not differ widely (Fielding and Well 1997). The computation of Kappa requires that the probabilistic output of the logostic regression procedure be dichotomized into predicted presences and absences, using a cut-off. When the true species distribution is known, this cut-off can be selected to maximize Kappa; but this is not possible if the models are applied to a new, uncensused area. In this situation, the modeler will have to choose between those cut-off values used by the person who created the models, or a subjective value - typically, 0.5. I computed Kappa using the cut-off points of the original models. However, because these values were tailored to the distribution of woodpeckers in the Jura, they may not be suited for the two Swiss sites. As an additional criterion of model performances, I mapped the predicted probabilities of presence in 25%-intervals. Visual comparisons with the true species distributions are an indication of how misleading the models may be, if the user did not know this true distribution and applied the models in lieu of census.

All statistical analyses were performed with SPSS v. 4.0.4 for the Macintosh (SPSS 1990).

Comparison of habitat variables

Although there was no IFN map for Switzerland, the Landsat TM scene overlapped part of the Jura study area (Fig. 3), providing an opportunity to compare the two datasets. For 353 cells of the Jura atlas, I obtained the composition of each IFN forest class in terms of Landsat TM classes. I used a paired t-test to compare edge density between the two data sets for the 353 cells, and to compare mean elevation for 197 cells at 50-m and 600-m resolution (the DEM was not available for the whole TM scene).

An important difference between the Landsat TM classification and the IFN map was the minimum mapping unit (MMU). The smallest polygon mapped by the IFN was about 4 ha (1 mm on the map), whereas MMU for the Landsat image was 0.0625 ha -- 64 times smaller. In an attempt to reduce differences between the two data sets, I generalized the Landsat image to increasingly larger MMUs, from 0.0625 ha to 1, 2, and 4 ha, using a merge program (Ma 1995). I computed basic statistics (number of patches, mean patch size, and percent of landscape) for each Landsat class at the four MMUs, and compared edge density between the original and each of the resampled image using a paired t-test.

RESULTS

Model performances

Model performances differed among study areas, species, and model type, but were generally poor (Table 2, Fig. 4a and 4b). Even when improvement over chance classification was statistically significant ($P \leq 0.05$), Kappa was never higher than

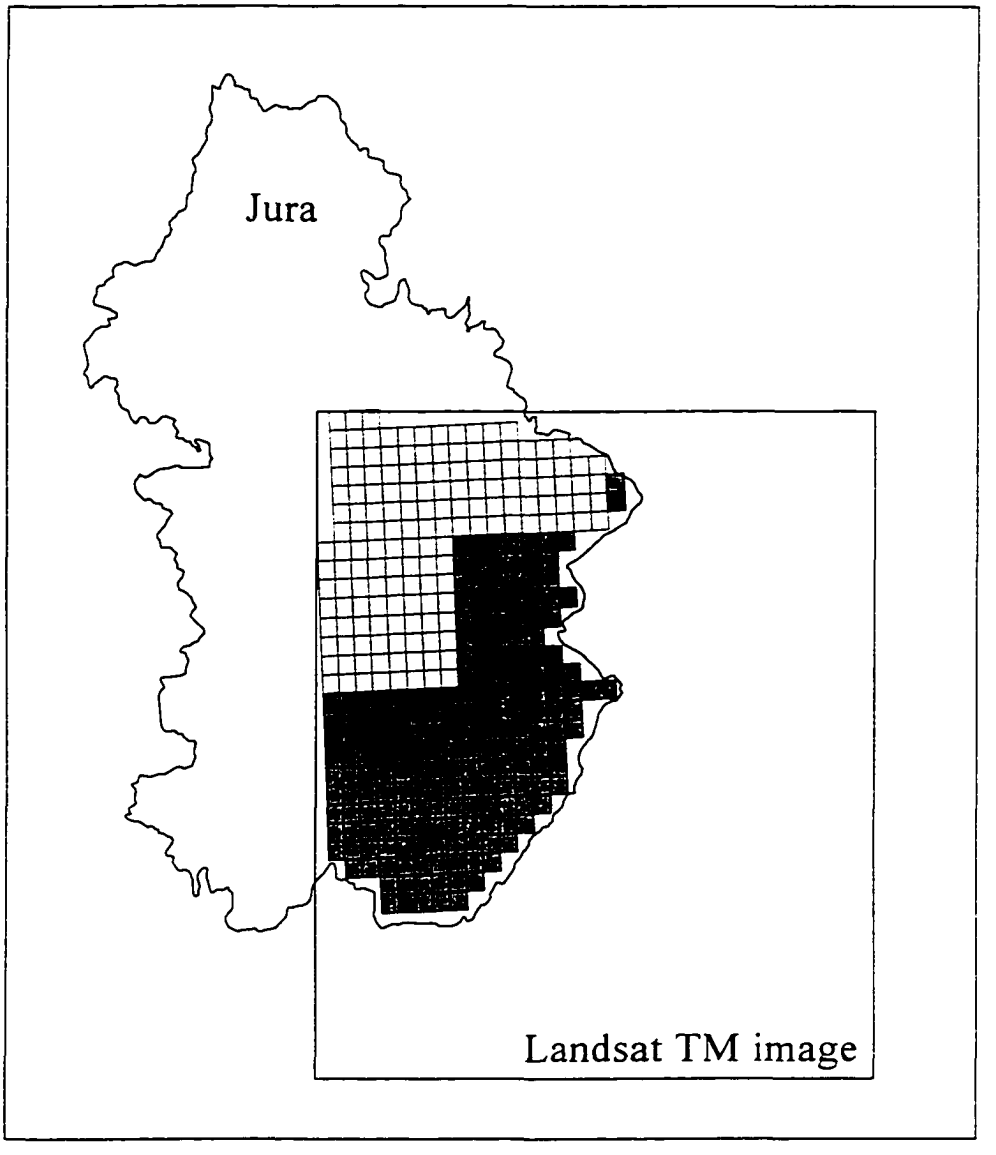


Figure 3. Census cells of the Breeding Bird Atlas of the Jura for which land cover composition data were available from both the Inventaire Forestier National map of the Jura, and a classified Landsat TM image. The darker shade of grey corresponds to cells for which elevation data could be extracted from a 50-m digital elevation model.

25%. The highest Kappa value was obtained for the universal model of the great spotted woodpecker at 2-ha MMU in the Orbe Valley ($K = 0.228$, $P < 0.0001$, Table 2). This model produces numerous errors of commission (absences with a probability of presence greater than 50%; Fig. 4a).

For species common to both study areas (the green and great spotted woodpeckers), models performed better in the Orbe Valley than in Geneva Canton (Table 2). Improvement over chance classification was never statistically significant for the black and grey-headed woodpeckers, nor for the wryneck; none of the presences of the middle spotted woodpecker was correctly classified. A cut-off of 0.5 would fail to identify any of the grey-headed, middle-spotted, and lesser spotted woodpeckers and wryneck presences in Geneva Canton (Fig. 4b).

Model type sometimes had a strong influence on model performances. There was no improvement over chance classification with the statistical model of the great spotted woodpecker, but K values were around 20% with the other two model types (Table 2). Similarly, K values were greater than 0 only with the statistical models of the grey-headed and lesser spotted woodpecker (Table 2).

Models in which edge density was entered as a variable, i.e., all three models for the green woodpecker, were most affected by MMU increase. In the Orbe Valley, K values almost doubled between 0.0625-ha MMU and 1-ha MMU, but either decreased or remained stable at higher MMUs (Table 2). In Geneva Canton, the only statistically significant improvement over chance classification for this species occurred for the biological model, at the 1-ha MMU ($K = 0.133$, $P < 0.025$, Table 2). There was no improvement over chance classification when the non-merged image was used, for all three model types, and K values were less than 10% for the statistical and universal models at all MMUs. K values also dropped below 10% for the biological model at

Table 2. Classification results (Cohen's Kappa) of three types of logistic regression models for seven woodpecker species at four different minimum mapping units in the Orbe Valley and in Geneva Canton, Switzerland. * = $P \leq 0.05$.

		Logistic regression model types											
		Biological				Parsimonious				Universal			
Species	Site	0.0625	1	2	4	0.0625	1	2	4	0.0625	1	2	4
Black	Orbe	0.029	0.029	0.029	0.029	0.045	0.047	0.035	0.034	0.022	0.028	0.029	0.028
Green	Orbe	0.106*	0.187*	0.120	0.131	0.068*	0.125*	0.107*	0.124*	0.082*	0.171*	0.168*	0.165*
	Geneva	0	0.133*	0.042	0.031	0	0.07	0.052	0.002	0	0.079	0.048	0
Grey-headed	Geneva	0	0	0	0	0.097	0.117	0.085	0.117	0	0	0	0
Great spotted	Orbe	0.190*	0.196*	0.196*	0.196*	0	0	0	0	0.212*	0.220*	0.228*	0.212*
	Geneva	0.076*	0.059	0.044	0.035	0.036	0.018	0.014	0.014	0.042	0.028	0.028	0.028
Middle spotted	Geneva	0	0	0	0	0	0	0	0	0	0	0	0
Lesser spotted	Geneva	0	0	0	0	0.150*	0.125*	0.112*	0.053	0	0	0	0
Wryneck	Geneva	0	0.010	0.018	0.021	0.004	0.019	0.016	0.019	0	0.012	0.002	0

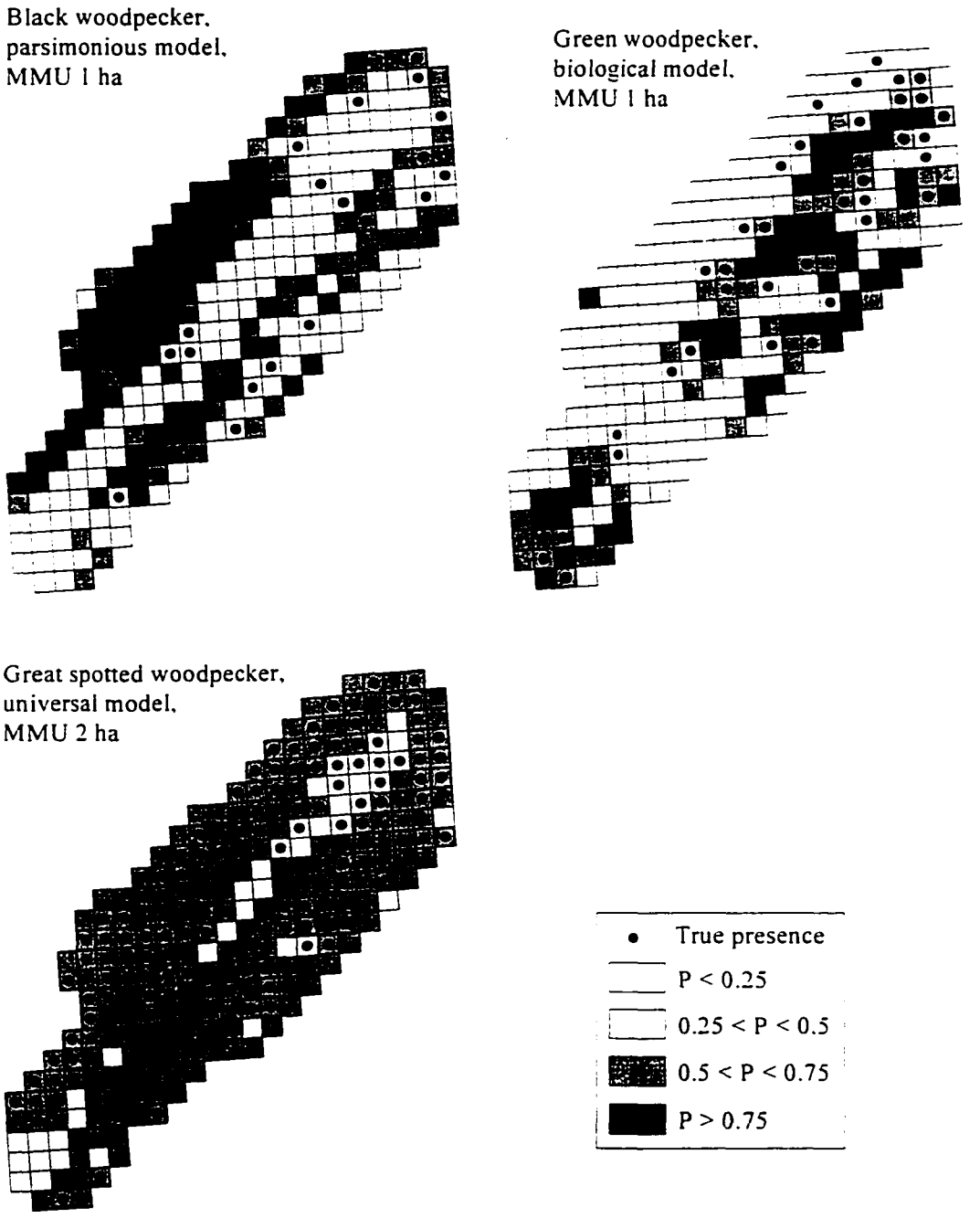


Figure 4a. Predicted probabilities of presence of woodpecker species in the Orbe Valley, Switzerland, obtained from logistic regression models developed in the Jura, France. Three model types and four different minimum mapping units (MMUs) were used; only the best model is presented for each species.

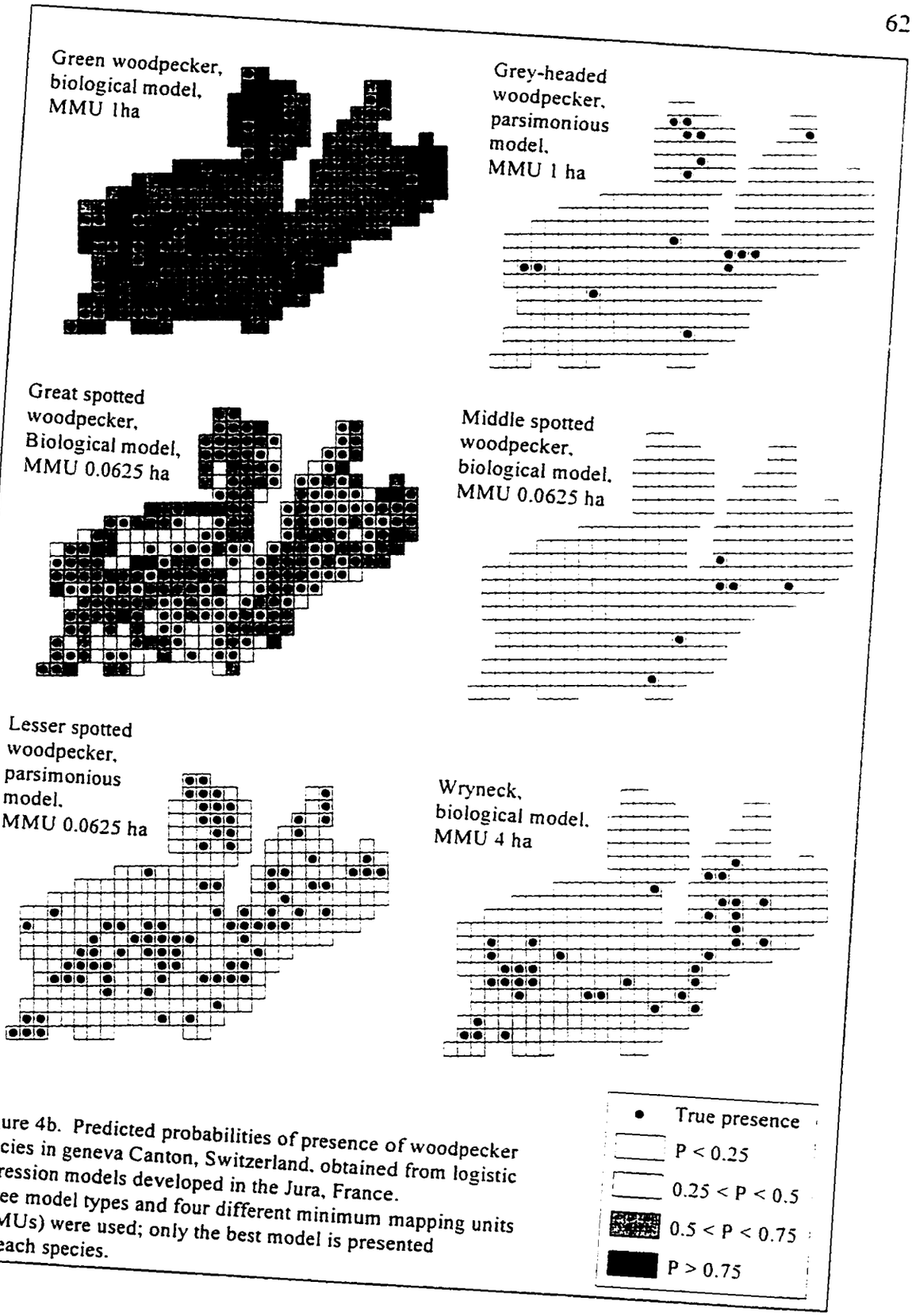


Figure 4b. Predicted probabilities of presence of woodpecker species in geneva Canton, Switzerland, obtained from logistic regression models developed in the Jura, France. Three model types and four different minimum mapping units (MMUs) were used; only the best model is presented for each species.

2-ha and 4-ha MMUs (Table 2).

Comparison of the original and test variables

Elevation did not differ significantly when computed from the 600 m or the 50 m data (paired t-test, $t = -1.25$, df = 196, 2-tailed $P = 0.212$). Edge density was much higher when computed from the Landsat classification (mean = 94.49, SD = 31.04) than from the IFN map (mean = 14.82, SD = 7.58), and the difference was highly significant (paired t-test, $t = -54.84$, df = 352, 2-tailed $P < 0.001$).

Land cover overlay of the IFN map and the Landsat classification revealed strong differences between the two classifications (Tables 3a and 3b). Highest correspondence was obtained between Open, Fir and Spruce in the IFN map, and Open and Mixed in the Landsat data (Table 3a). No single Landsat class dominated the composition of the other IFN classes; for example, the IFN class Beech was only 40% Beech; Xeric forests, only 30% Xeric; OtherCut was Deciduous, Beech, Mixed, Planted, and Open, in similar proportions (Table 3a). Grouping the classes confirmed the similarity of Mixed and Open, and the relatively poor correspondence of the other classes (Table 3b).

Landscape configuration (i.e., the spatial arrangement of patches), more than composition, was profoundly affected when the MMU of the Landsat TM classification was increased from 0.0625 ha to 4 ha, with the biggest change occurring between 0.0625 ha and 1 ha (Tables 4a and 4b, Fig. 5). In both study areas, number of patches decreased and mean patch size increased. Percent of the landscape in the different classes changed more for the Orbe Valley (Table 4a) than for Geneva Canton (Table 4b). In the Orbe Valley, small clumps of Beech and Plantation pixels were converted to Mixed (Fig. 3), but the ratio of forested to open lands remained constant (about two

Table 3a. Composition (%) of 13 cover types from the Inventaire Forestier National (IFN) map of the Jura, France, when overlaid to a classified Landsat TM image for 353 census cells of the Breeding Bird Atlas of the Jura, France.

	Percent composition of Landsat TM classes						
IFN classes	Xeric	Deciduous	Beech	Mixed	Planted	Open	Total
MatDecid	0	44.21	26.90	2.53	1.50	24.87	100
Ceppice	0	48.04	20.10	3.88	5.57	22.40	100
Xeric	29.55	10.71	24.64	5.67	2.49	26.94	100
Scree	9.5	13.82	34.71	14.72	6.08	21.16	100
MtnDecid	10.6	39.76	24.72	7.62	4.70	12.62	100
Beech	6.35	10.10	40.08	21.73	5.16	16.58	100
Fir	1.86	1.81	17.13	64.68	8.51	6.01	100
Spruce	0.44	0.26	14.77	69.11	5.10	10.31	100
ShelterCut	1.97	51.36	17.70	9.42	11.20	8.35	100
OtherCut	4.24	16.95	18.29	26.85	19.35	14.32	100
Patchy	5.34	9.61	22.84	21.00	9.00	32.22	100
RipHetero	12.25	14.35	18.38	6.49	3.20	45.31	100
Open	2.41	3.52	7.78	4.01	1.66	80.63	100

Table 3b. Composition (%) of cover types from the Inventaire Forestier National (IFN) map of the Jura, France, when overlaid to a classified Landsat TM image for 353 cells of the Breeding Bird Atlas of the Jura, France. The original 13 cover types of the IFN map have been regrouped into five general classes, and the six classes of the Landsat image into four.

Regrouped IFN classes	Percent composition of regrouped Landsat TM classes				
	Decid	Mixed	Planted	Open	Total
Decid	45.82	32.64	4.01	17.83	100
Mixed	5.49	78.17	6.74	9.60	100
Planted	36.94	36.30	15.35	11.40	100
Patchy	22.91	30.88	5.04	41.17	100
Open	5.93	11.79	1.66	80.63	100

Table 4a. Effect of changing the minimum mapping unit (MMU) on patches of various cover types in the Orbe Valley test site, Switzerland.

	MMU (ha)	Number of patches	Mean patch size (ha)	Percent of landscape
Beech	0.0625	18081	0.19	10.59
	1	394	4.04	4.94
	2	169	7.83	4.11
	4	80	13.35	3.32
Mixed	0.0625	5155	3.12	49.94
	1	237	79.54	58.57
	2	115	166.32	59.43
	4	59	329.25	60.36
Planted	0.0625	9685	0.09	2.68
	1	6	1.98	0.04
	2	1	5.88	0.02
	4	1	5.88	0.02
Open	0.0625	9393	1.26	36.79
	1	284	41.30	36.45
	2	169	69.40	36.44
	4	107	109.21	36.31

Table 4b. Effect of changing the minimum mapping unit (MMU) on patches of various cover types in the Geneva Canton test site, Switzerland.

	MMU (ha)	Number of patches	Mean patch size (ha)	Percent of landscape
Decid	0.0625	11998	0.25	8.19
	1	329	7.26	6.40
	2	166	13.16	5.86
	4	92	22.21	5.48
Beech	0.0625	12711	0.16	5.35
	1	171	4.44	2.04
	2	85	8.04	1.83
	4	38	15.79	1.61
Mixed	0.0625	4390	0.14	1.59
	1	36	4.94	0.48
	2	16	9.61	0.41
	4	11		0.38
Planted	0.0625	3074	0.09	0.74
	1	23	2.91	0.18
	2	7	4.52	0.08
	4	1	12.98	0.01
Open	0.0625	3358	9.34	84.13
	1	47	721.39	90.91
	2	18	1902.44	91.81
	4	7	4929.58	92.52

Table 5. Effect of increasing the minimum mapping unit (MMU) on edge density values (m/ha) in 100-ha cells in the Orbe Valley and Geneva Canton, Switzerland.

	Orbe Valley		Geneva Canton	
MMU	Mean	Stddev	Mean	Stddev
0.0625	106.85	49.34	83.88	48.722
1	47.49	31.73	17.48	23.06
2	42.07	29.03	13.33	20.34
4	37.90	27.99	10.86	19.17

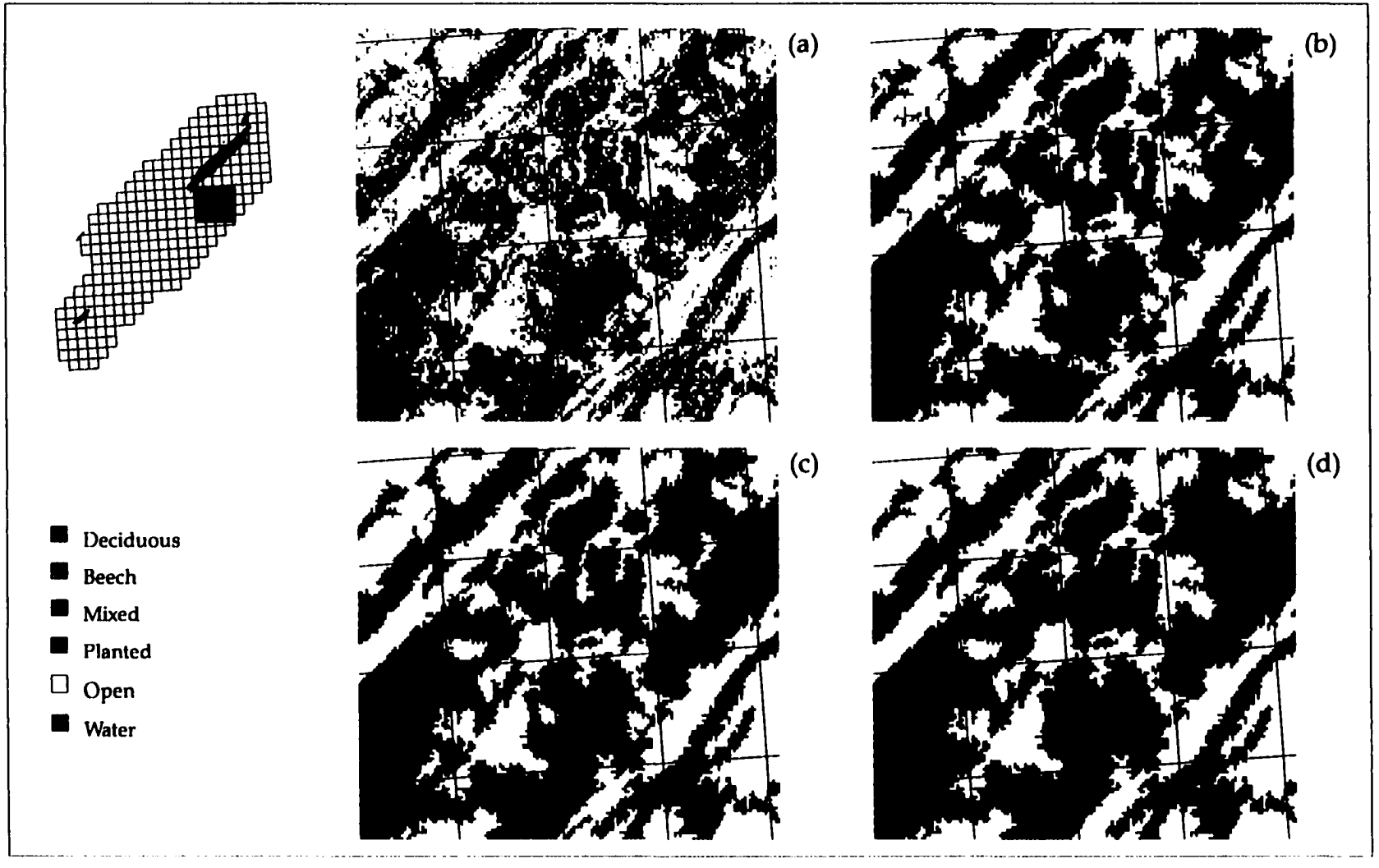


Figure 5. Effect of changing the minimum mapping unit (MMU) on the land cover map of the Orbe Valley, Switzerland. (a) MMU 0.0625 ha; (b) MMU 1 ha; (c) MMU 2 ha; (d) MMU 4 ha.

thirds forested for one third open). In Geneva Canton, all forest classes lost some pixels to open areas, which increased from 84.13% (0.0625-ha MMU) to 90.91% (1-ha MMU) of the landscape (Table 4b).

A direct consequence of the elimination of isolated pixels and small patches, edge density decreased sharply as MMU increased, with again a major step between 0.0625-ha MMU and 1-ha MMU (Table 5). In Geneva Canton, edge density values of the merged images were similar to those of the IFN map, but they remained about three times larger in the Orbe Valley.

DISCUSSION

I tested the universality of woodpecker-habitat relationship models developed in the Jura by applying them to two sites in Switzerland, where the source for habitat variables differed from the source of the variables used in model development. Model performances, as measured by Cohen's Kappa and visual comparison of predicted vs true presences, were low, regardless of species, study site, and minimum mapping unit of the land cover map. The highest improvement over chance classification (22.8%) was well below the 40% suggested by Landis and Koch (1977) as a minimum for good reliability. Several factors may have contributed to the relatively poor predictive ability of the models: 1) inappropriateness of the independent variables used in the testing; 2) poor classification rates of the original models; 3) difference in the scale used for model development and model testing; 4) differences among breeding bird atlases; 5) inappropriateness of the modeling approach (logistic regression); and 6) geographic variations in habitat occupancy.

Variables used in the testing

Important differences between predictor variables used for model development and model testing are likely to adversely affect predictive reliability. Overlay procedures identified such differences between the IFN forest cover map and the Landsat classification. Several factors may account for these differences, such as different MMUs and classification schemes. The MMU of the Landsat image was a 25-m pixel (i.e., 0.0625 ha), 64 times smaller than that of the IFN map (about 4 ha); patches smaller than 4 ha did not show up on the IFN map, but appeared on the Landsat image as single pixels or small groups of pixels. For example, small forest openings were visible on the Landsat image where the IFN map displayed a continuous forest patch; this may explain why the Landsat class Open composed between 6% and 45% of the 12 IFN forest classes (Table 3a). A large proportion (45%) of Open pixels in the class RipHetero is quite normal, because this class was defined as small clumps of trees connected by a network of edges or wooded pasture (Chapter 1). Here certainly lies the main source of difference between the two maps: the Landsat image classification was based on canopy reflectance values, regardless of forest management, whereas forestry practices formed the basis of the definition of the IFN classes. For example, the IFN class Shelterwood Cuts consists of conifers (usually fir, *Abies alba*) "hidden" under an overstory of mature deciduous trees (IFN 1980, pers. obs.); in the Landsat TM classification, these patches showed as being predominantly Deciduous and Beech (Table 3a). Other IFN classes were differentiated on the basis of stand structure and management. MatDecid, Coppice, and MtnDecid all showed about the same proportion of oak (*Quercus* sp.) and beech (*Fagus sylvatica*) trees (Table 3a), but the age structure of the stands differs (IFN 1980, pers. obs.). Stand characteristics other than composition did not influence the Landsat classification scheme, so all three classes

blend (based on the proportions of Decid and Beech; Table 3a). The IFN class Patchy represents an extreme case, because tree species composition was not even part of the class definition (Chapter 1).

Changing the minimum mapping unit of the Landsat TM classification to make it more similar to that of the IFN map had only minimal effect on model performances, because only edge density was significantly affected. Increasing the MMU improved classification success of the green woodpecker models; in Geneva Canton, however, improvement over chance classification was lower at 2 ha and 4 ha MMUs than at 1 ha MMUs, possibly because at these MMUs, areas of suitable habitat (small patches of trees) were lost from the map.

Also of concern is the quality of the Landsat TM classification. A portion of the TM scene was intensively ground-truthed and classified using field data (Vuillod 1994). The resulting classification algorithm was then used to classify the rest of the scene, but no accuracy assessment was conducted (Vuillod 1995). Classification errors are certainly present; for example, isolated pixels of Open may be real (small forest openings), but the numerous isolated pixels of Planted scattered across both test sites (See Fig. 3 for an example) are probably artifacts of the classification method, because forest management practices in the Jura mountains are most often applied to areas larger than a single pixel. These errors, however, probably contributed little to model failure, compared to the important differences between data sets presented above.

Classification ability of the original models

Model performances in the Jura varied more among species than among model types (Chapter 1). Best results were obtained for the black and middle spotted woodpeckers (Kappa > 30%), but even for these two species, there were numerous

errors of both commission (absences predicted as presences) and omission (presences not predicted). Models with low classification success tend to be poor predictors when applied to different areas (Fielding and Haworth 1995).

Surprisingly, the universal model for the great spotted woodpecker performed better in the Orbe Valley ($0.212 \leq K \leq 0.228$; Table 2) than it did in the Jura ($K = 0.160$; Chapter 1). A possible explanation is that because the model emphasized the negative association between the species' presence and open areas (Chapter 1), it was well suited to "dichotomous" nature of the Orbe Valley landscape (with prairies concentrated along the valley's axis, and large forest patches covering the slopes). The Jura, by contrast, comprised a variety of landscapes (Chapter 1), and although fewer presences were censused in forested squares than in non-forested ones, the model was too simplistic to perform well there.

Scale issues

The area over which models were developed covered about 500,000 ha and included a variety of landscapes, ranging from lowlands rich in deciduous forests and open, agricultural lands, to the dense coniferous forests of the Upper Jura (Chapter 1). The models were tested, however, on smaller and more homogeneous sites in Switzerland. The Orbe Valley test site covered 27,300 ha and was more similar to the Upper Jura region, whereas Geneva Canton consisted of 30,600 ha dominated by crops, deciduous forests, the Lake of Geneva, and urban development - overall, a landscape closer to that of the Plains region of the Jura. Hence, the test sites appear to be reasonable subsets of the area used for model development.

Scale is an important consideration in habitat selection studies, as patterns of wildlife-habitat associations tend to change from one scale to another (Wiens 1989).

Therefore, models developed at one spatial scale may not work well when applied to a different one. This explains why the parsimonious model of the great spotted woodpecker failed completely in the Orbe Valley (no improvement over chance classification; Table 2). At the broader spatial scale of the Jura, species' presence was negatively correlated with elevation (Mann-Whitney U-test, $P \leq 0.001$), even though mean elevation of presence cells ranged from 190 m to 1287 m (Chapter 1, Table 3). Elevation negatively entered the parsimonious model, which therefore failed to correctly classify presences when applied to the high-elevation Orbe Valley (mean elevation 1242 m, range 993 m - 1585 m). Classification accuracy increased when elevation did not enter the models (Table 2), and in Geneva Canton (mean elevation 423 m) at least some of the presences were correctly classified ($K > 0$; Table 2).

As an additional scale issue, it is important to note that the models were developed using 575-ha cells as the units of analysis, but were tested on 100-ha cells. Wiens (1985) and Wiens *et al.* (1987), in a study of the patterns of habitat occupancy of North American shrubsteppe birds, found that patterns of bird-habitat associations may reverse from one scale to another.

Breeding bird atlases

All three breeding bird atlases (Jura, Orbe Valley, Geneva Canton) were created from census data collected by local ornithologists and birders. The methods used to collect the data may have varied from one site to another; particularly, the smaller census scale of the Swiss atlases may have permitted a more thorough coverage of each cell by the observers, hence making it less likely that a woodpecker nest would be missed. An overlap between the Jura and the Orbe Valley altases, for 7 Jura cells, made it possible to compare census results between the two atlases. For the black woodpecker, census

results were in agreement for 6 cells; one Jura cell coded as presence did not contain any presence in the Orbe atlas. The same proportions (6 agreements, one disagreement) were found for the green and great spotted woodpecker woodpecker, but this time a presence was noted in several of the smaller Orbe cells but none in the overlapping Jura cell. These differences could also be due to temporal factors: data were collected from 1986 to 1992 for the Jura atlas, but from 1980 to 1982 for the Orbe atlas.

Modeling approach

The structure of logistic regression models, or any other type of mathematical model using specific coefficients, may limit their universality (Fielding and Haworth 1995). Statistical methods "force" the data to fit a particular model, and such algorithmic solutions may impose unreal structure to the data (Fielding 1994). Other approaches, such as artificial neural networks or genetic algorithms, avoid the application of constraining rules (Fielding 1994), but have not yet been applied to wildlife-habitat studies.

Variations in habitat occupancy

Species-habitat associations may vary geographically (e.g. Collins 1983, Shy 1984). This variation is usually more important at fine scales than at broader ones, where local differences in habitat selection are masked (Wiens 1989). The models were built using broad-scale variables over a reasonably large area (5000 km^2), but regional differences in nesting habitat characteristics may still occur between the Jura and the two Swiss test sites, despite their close geographic proximity (Fig. 1). For example, the middle spotted woodpecker is traditionally associated with large forests of mature oak and hornbeam (*Carpinus betulus*) (Short 1982, Wesolowski and Tomialojc 1986,

Spitznagel 1990, Jovéniaux 1993, Angelstam and Mikusinski 1994). The models developed in Chapter 1 reflect the importance of mature deciduous forests to the species. Half of the presences were located in one very large oak forest, supporting the suggestion that the species prefers woods of 30 ha or more (Cramp 1985). In Switzerland, however, use of old orchards and small patches of lime trees (*Tilia* sp.) has been documented (Jovéniaux 1993). Although not unheard of (Cramp 1985, Clergeau and Chefson 1988), this choice of nesting habitat is uncommon and has not been documented in the Jura (Jovéniaux 1993). Therefore, differences in nesting habitat preferences between the two areas are likely to have contributed to the total failure of the models for the middle spotted woodpecker in Geneva Canton (no absence correctly predicted; Table 2).

CONCLUSION

Very rarely (if ever) do models perform perfectly. The intended use of a model will influence the amount of incorrect prediction that is acceptable (Salwasser 1986), and the ability of a model to meet its purpose is the ultimate criterion by which the model should be evaluated (Starfield 1997). Woodpecker-habitat models developed in the Jura failed to predict the distribution of these species in two areas in neighboring Switzerland, even if low accuracy standards are used: in many instances, the models did not provide improvement over chance classification. Important differences between the IFN map and the Landsat image were largely responsible for these poor results. Testing this suggestion, however, would require that the models be run twice, once with an IFN map as a source of data, once with the Landsat image, so that the relative role of using a different data set could be detangled from other factors (such as those presented above).

Ideally, the test data should also have been collected at the same spatial scale (i.e., same atlas cell size) and temporal scale (bird censuses conducted at the same period) as those of the Jura atlas. I could not pursue these alternatives, given my dependence on existing data. For models developed with the goal of prediction, efforts should be made to minimize differences between the datasets from which habitat variables are extracted. For example, Gates *et al.*(1995) estimated that changes in the recording method of the agricultural census data that entered their bird-habitat models may have contributed to the inability of the 1988-based models to predict species' distributions in 1969. Birchman and Jordan (1996) questioned the legitimity of comparing the distribution maps from two breeding bird atlases (1976 and 1993) as a mean to study changes in distribution, because of differences in methodology between the two atlases (but see Greenwood *et al.* 1996). Classified satellite imagery, a suitable source of habitat variables for regional-scale studies (Palmeirin 1988, Shaw and Atkinson 1990), can be used to minimize differences between land-cover data used for model development and testing. However, extracting habitat variables from the same source does not guarantee good model performances during testing (see Fielding and Haworth 1995; Chapter 3).

REFERENCES

Angelstam, P, and G. Mikusinski. 1994. Woodpecker assemblages in natural and managed boreal and hemiboreal forest - a review. Ann. Zool. Fennici **31**. 157-172.

Bircham, P. M. M, and W. J. Jordan. 1996. A consideration of some of the changes in distribution of "common birds" as revealed by *The New Atlas of Breeding Birds in Britain and Ireland*. Ibis **139**, 183-186.

Capen, D. E., ed. 1981. The use of multivariate statistics in studies of wildlife habitat. USDA For. Serv. Gen. Tech. Rep. RM-87, Rocky Mtn. Forest and Range Exp. Sta., Fort Collins, CO.

Capen, D. E., J. W. Fenwick, D. B. Inkley, and A. C. Boynton. 1986. Multivariate models of songbird habitat in New England forests. Pages 171-175 *in* J. Verner, M. L. Morrison, and C. J. Ralph, eds. Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates. Univ. of Wisconsin Press, Madison.

Chalk, D. E. 1986. Summary: development, testing, and application of wildlife-habitat models - the researcher's viewpoint. Pages 155-156 *in* J. Verner, M. L. Morrison, and C. J. Ralph, eds. Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates. Univ. of Wisconsin Press, Madison.

Chandler, S. K., J. D. Fraser, D. A. Buehler, and J. K. Seegar. 1995. Perch trees and shoreline development as predictors of bald eagle distribution on Chesapeake Bay. J. Wildl. Manage. **59**, 325-332.

Clergeau, P., and P. Chefson. 1988. Les Pics. Série "Comment vivent-ils" Vol. 20. Atlas Visuels Payot Lausanne. Payot Lausanne, Lausanne.

Collins, S. L. 1983. Geographic variation in habitat structure of the black-throated green warbler (*Dendroica virens*). Auk **100**, 382-389.

Cramp, S. 1985. The Birds of Western Palaearctic, Vol. IV. Oxford Univ. Press, Oxford.

ESRI. 1995. Arc/Info 7.0.3. Redlands, CA.

Fielding, A. H. 1994. A review of methods used to investigate bird-habitat associations. SNH review no. 4. Scottish Natural Heritage, Edinburgh.

Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conserv. Biol. **9**, 1466-1481.

Gates, S., D. W. Gibbons, P. C. Lack, and R. J. Fuller. 1993. Declining farmland bird species: modelling geographical patterns of abundance in Britain. Pages 153-177 *in* P. J. Edwards, R. M. May, and N. R. Webb, eds. Large-Scale Ecology and Conservation Biology. Blackwell Sci. Publ., London.

Géroudet, P., C. Guex, M. Maire, and others. 1983. Les Oiseaux Nicheurs du Canton de Genève. Muséum de Genève, Suisse.

Glayre, D., and D. Magnenat. 1984. Oiseaux Nicheurs de la Haute Vallée de l'Orbe. Nos Oiseaux 398. Fascicule spécial du volume 37.

Greenwood, J. J. D., R. J. Fuller, and D. W. Gibbons. 1996. Assessing changes in distribution from atlas data: reply to Birchman and Jordan. Ibis **139**, 186-189.

IFN 1980. Département du Jura. Résultats du Deuxième Inventaire Forestier. Tome 1. Ministère de l'Agriculture, Direction des Forêts.

IFN 1993. Carte Forestière du département du Jura. Ministère de l'Agriculture, Direction des Forêts.

Jovéniaux, A. 1993. Atlas des Oiseaux Nicheurs du Jura. G.O.J. ed., France.

Landis, J. R., and G. C. Koch. 1977. The measurement of observer agreement for

categorical data. Biometrics **33**, 159-174.
Ma, Z. 1995. Using a rule-based merging algorithm to eliminate 'salt/pepper' and small regions of classified image. Ninth Annual Symposium on Geographic Information Systems, Vancouver, B.C., Canada, 834-837.
Marcot, B. G., M. G. Raphael, and K. H. Berry. 1983. Monitoring wildlife habitat and validation of wildlife-habitat relationships models. Proc. North Am. Wildl. Conf. **48**, 315-329.
Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1992. Wildlife-Habitat Relationships. Univ. of Wisconsin Press, Madison.
Nadeau, S., R. Décarie, D. Lambert, and M. St-Georges. 1995. Nonlinear modeling of muskrat use of habitat. J. Wildl. Manage. **59**, 110-117.
Noon, B.R. Biometric approaches to modeling - the researcher's viewpoint. Pages 197-201 *in* J. Verner, M. L. Morrison, and C. J. Ralph, eds. Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates. Univ. of Wisconsin Press, Madison.
Palmeirin, J. M. 1988. Automatic mapping of avian species habitat using satellite imagery. Oikos, **52**, 59-68.
Salwasser, H. 1986. Modeling habitat relationships of terrestrial vertebrates - the manager's viewpoint. Pages 419-429 *in* J. Verner, M. L. Morrison, and C. J. Ralph, eds. Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates. Univ. of Wisconsin Press, Madison.
Shaw, D. M., and S. F. Atkinson. 1992. An introduction to the use of geographic information systems for ornithological research. Condor **92**, 564-570.
Short, L. L. 1982. Woodpeckers of the World. Delaware Museum of Natural History Monograph Series No. 4. Delaware Museum of Natural History, Greenville.
Shugart, H. H., Jr. 1981. An overview of multivariate methods and their application to studies of wildlife habitat. Pages 4-10 *in* D. E. Capen, ed. The use of multivariate statistics in studies of wildlife habitat. USDA For. Serv. Gen. Tech. Rep. RM-87. Rocky Mtn. Forest and Range Exp. Sta., Fort Collins, CO.
Shy, E. 1984. Habitat shift and geographic variation in North American tanagers (Thraupinea: *Piranga*). Oecologia **63**, 281-285.
Spitznagel, A. 1990. The influence of forest management on woodpecker density and habitat use in floodplain forests of the Upper Rhine Valley. Pages 117-146 *in* A. Carlson and G. Aulén, eds. Conservation and Management of Woodpecker Populations. Report 17, Swedish Univ. of Agric. Sci., Uppsala.
SPSS, Inc. 1990. SPSS version 4.0.4. for the Macintosh. Chicago, IL, U.S.A.
Starfield, A. M. 1997. A pragmatic approach to modeling for wildlife management. J. Wildl. Manage. **61**, 261-270.
Titus, K., J. A. Mosher, and B. K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: ecological applications. Am. Midl. Nat. **111**, 1-7.
Verbyla, D. L., and J. A. Litvaitis. 1989. Resampling methods for evaluating classification accuracy of wildlife habitat models. Environ. Manage. **13**, 783-787.
Verner, J., M. L. Morrison, and C. J. Ralph, eds. 1986. Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates. Univ. of Wisconsin Press, Madison.
Vuillod, P. 1994. Paysage visible et aménagement: modélisations cartographiques et test sur le Haut-Jura. Ph.D. dissertation, Université de Franche-Comté, France.

Vuillod, P. 1995. Classification d'une image satellite Landsat pour le Parc Naturel Régional du Haut-Jura. Unpubl. Report.

Wesolowski, T., and L. Tomialojc. 1986. The breeding ecology of woodpeckers in a temperate primaeval forest - preliminary data. Acta Ornithologica **22**. 1-21.

Wiens, J. A. 1985. Habitat selection in variable environments: shrub-steppe birds. Pages 227-251 *in* M. L. Cody, ed. Habitat Selection in Birds. Academic Press, Inc.

Wiens, J. A. 1989. The ecology of bird communities. Vol. 1. Foundations and patterns. Cambridge Univ. Press.

Wiens, J. A., J. T. Rottenberry, and B. Van Horne. 1987. Habitat occupancy patterns of North American shrubsteppe birds: the effects of spatial scale. Oikos **48**. 132-147.

CHAPTER 3

ASSESSING THE UNIVERSALITY OF WILDLIFE-HABITAT RELATIONSHIP MODELS: AN EXAMPLE WITH THE GREEN WOODPECKER (PICUS VIRIDIS) IN SWITZERLAND

Key words: habitat modeling, *Picus viridis*, model universality, scale, Switzerland.

ABSTRACT

I used logistic regression to relate the presence of the green woodpecker (*Picus viridis*) to simple, broad-scale habitat variables in two study sites in Switzerland, to assess the effect of scale on classification results and universality of bird-habitat models. Scale changes were made by enlarging the mesh size of the breeding bird atlas maps (from 100 ha to 400 ha), and increasing the minimum mapping unit (MMU) of the classified Landsat image from which land cover composition was extracted (from 0.0625 ha to 1 ha). Models developed in one site were applied to data in the other. Model composition differed between the two sites, suggesting that the predictor variables were not appropriate for developing general models, because they did not reflect differences in landscape configuration between the sites. Model composition, classification results, and model universality were all affected by scale, and the changes were site-specific. This variability suggests that caution should be used when applying a model developed in one part of a species' range to another, and that particular attention should be given to the choice of the scale of analysis.

INTRODUCTION

Scale influences every aspect of ecological research (Wiens 1989, Levin 1992), and models relating the distribution of wildlife species to characteristics of their environment are no exception. Measures of habitat variables may be influenced by the scale at which they are collected. For example, the characteristics of a landscape, i.e., the relative proportion of land cover types and their arrangement in patches, are scale-dependent: increasing the minimum mapping unit (MMU) of a raster map will affect landscape composition and configuration through the loss of certain cover types (Turner *et al.* 1989). These changes will, in turn, affect modeling output. For example, Stoms (1992) showed that the distribution of "hot spots" of vertebrate species richness, obtained from predicted species distribution maps, varied spatially as MMU increased, because the generalization of the raster habitat map caused some habitat types to be locally eliminated and fewer species to be predicted. Simplifying the habitat map may affect the predictive ability of a model in one of two ways: classification success increases, because the "noise" that obscured patterns is eliminated; or, classification success decreases, because small but important habitat patches associated with species' presence are gone.

Models are also likely to be affected by the scale of the modeled species' distribution data. Fine-scale censuses collect information in the form of points, such as nest locations; but for large areas, distribution data are more commonly presented as atlas maps in a grid format. The scale of such grids varies, from 1 or 2 km for local atlases (e.g., Glayre and Magnenat 1984, Thomas and Abery 1995), to 30 km for entire countries (e.g. Robertson *et al.* 1994). Again, two outcomes are possible: a broader census scale may either reveal patterns hidden by individual variability (as with a species

that nests in a variety of tree species, but is always found in a certain forest type), or fail to identify essential fine-scale features (for example, the great spotted woodpecker *Dendrocopos major* will nest in many forest types, as long as snags are present).

In addition to influencing the classification success of a model, scale may affect its universality. Predictive modeling, the estimation of distribution or abundance of a wildlife species given information on habitat conditions (Morrison *et al.* 1992), is an important tool for wildlife habitat management and conservation. Predictive models can be used for at least three types of applications: 1) to map species' distributions in areas that cannot be reached easily and/or censused (e.g., Osborne and Tigar 1992), or when distribution data are lacking (e.g., Gap Analysis; Scott *et al.* 1993); 2) to estimate whether an area is suitable for a given species (e.g., habitat suitability index models; Schamberger *et al.* 1982); and 3) to assess how species would respond to modifications of their habitat, such as those resulting from climate changes, land-use changes, or from alternative management scenarios. The success of these modeling efforts will be a function of model universality, i.e., a model's applicability to circumstances different from those under which it was developed. Model universality depends on the strength of the associations between a given species and its habitat, and on the appropriateness of the scale at which the model was developed (Fielding and Haworth 1995). For example, Collins (1983) showed that habitat structure and composition of the black-throated green warbler (*Dendroica virens*) varied across the species' range. The absence of consistent bird-habitat association patterns for that species would result in poor model performances if a model developed using distribution data from one part of the species' range was used to predict the species' distribution in another part of the range. Similarly, Converse and Morzuch (1981) found that the sign of correlations between snowshoe hare presence (*Lepus americanus*) and several habitat variables

reversed in different parts of the species' range. In such a case, increasing the scale of analysis may improve model generality by uncovering broad-scale patterns of species-habitat associations, but it may also mask real differences among areas, with unfortunate consequences if the models are used for management (Converse and Morzuch 1981). Also, a search for generality may not be possible if variability in the dependent variable cannot consistently be captured by the predictor variables, and systems are inherently unpredictable (Judson 1994).

In Europe, the long history of human disturbance to the environment has lead to the creation of a variety of landscapes, ranging from semi-natural to highly modified, depending on local history, topography, climate, and vegetation. Along this gradient, the Upper Orbe Valley, Switzerland (hereafter Orbe Valley) would be considered more "natural", in contrast to the more intensively developed, neighboring Geneva Canton. Breeding bird atlases exist for both areas (Géroudet *et al.* 1983 , Glayre and Magnenat 1984). Some species, such as the black or middle spotted woodpeckers, occur in only one of the two landscapes; others, such as the green woodpecker, have been censused in both. Because the green woodpecker is an edge species (Short 1982, Cramp 1985, Clergeau and Chefson 1988), and edge is a function of the MMU of a map (Chapter 2), models relating its presence to environmental variables are especially likely to be sensitive to scale. In this chapter, I use atlas distribution maps of the green woodpecker from the two Swiss atlases to: 1) determine if the associations between the green woodpecker's presence and simple, broad-scale habitat variables are strong and consistent enough to allow model universality; and 2) assess the effect of changing the scale of distribution data (atlas cell size) and that of habitat variables (MMU) on model classification results and universality.

STUDY AREAS

The Orbe Valley and Geneva Canton are both situated in the western part of Switzerland, although a portion of the Orbe Valley extends into eastern France (Fig. 1). The Orbe Valley is characteristic of high elevation valleys of the Jura mountain range (PNRHJ 1988): the valley floor is open pastures, surrounded by dense, unbroken mixed forests dominated by spruce (*Picea abies*) and beech (*Fagus sylvatica*). Urban development is minimal (less than 1%; Table 1). Elevation ranges from 972 m to 1669 m. Forestry, dairy farming, and tourism are the principal economic activities. By contrast, Geneva Canton is a highly developed agricultural landscape dominated by crops and fields (Table 1), with important urban and aquatic components (the city of Geneva and the Lake of Geneva; Table 1). Forests, mostly deciduous, occur as small patches embedded in the agricultural matrix, and elevation ranges from 328 m to 563 m.

METHODS

Digital database

Both atlases present green woodpecker breeding distribution data in the form of 1-km^2 (i.e., 100 ha) grid cells (Géroudet and Guex 1983, Glayre and Magnenat 1984). Green woodpecker breeding was labelled as "possible" or "certain" for 49 of the 273 cells of the Orbe Valley, and for 203 of the 306 cells of Geneva Canton. I used the Geographic Information System (GIS) software Arc/Info v. 7.0.3. (ESRI 1995) on a Unix workstation to digitally recreate the atlas grids, and simplified the coding by labelling both possible and certain breeding as "presence".

Data availability constrained the selection of habitat variables entering the models.

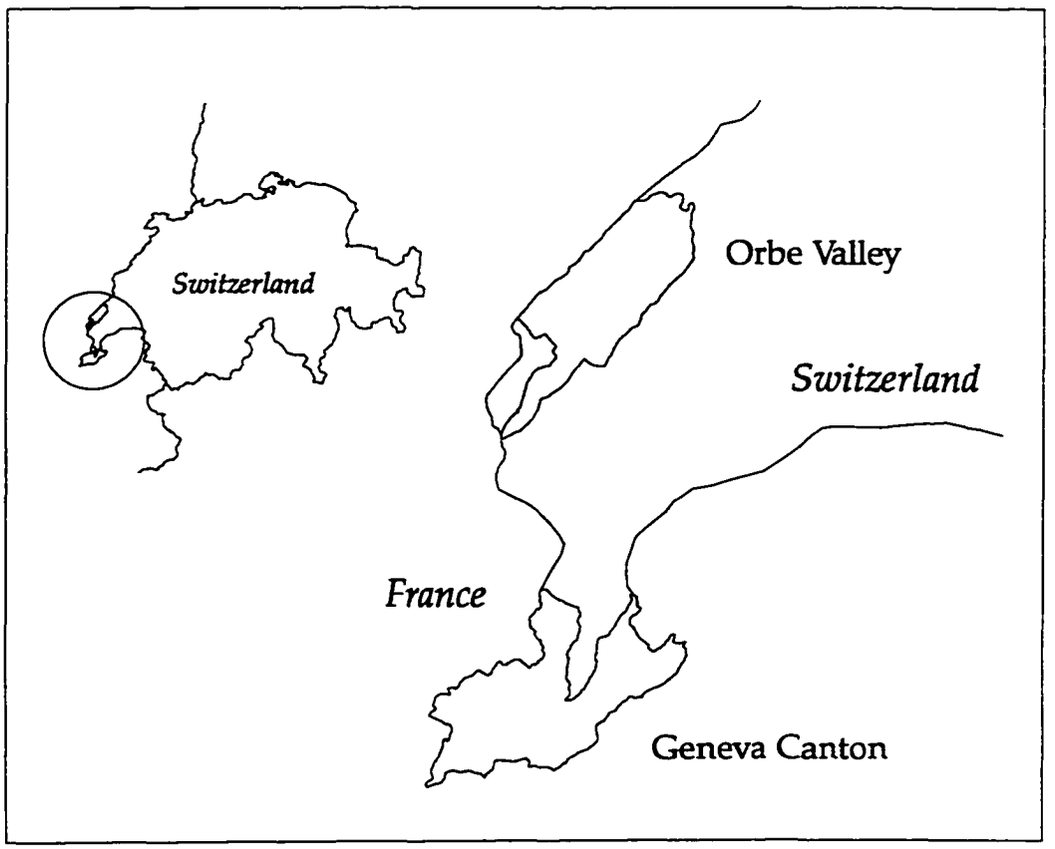


Figure 1. Location of the two study areas.

Table 1. Land cover composition of the Orbe Valley and Geneva Canton, Switzerland, as obtained from a classified Landsat TM satellite image.

	Orbe Valley		Geneva Canton	
Land cover class	Area (ha)	Percent	Area (ha)	Percent
Oak-hornbeam-beech forests	0.81	0.00	3055.13	8.19
Beech forests	3409.19	10.59	1994.69	5.35
Beech-fir-spruce forests	16071.44	49.94	594.44	1.59
Conifer plantations	861.31	2.68	275.94	0.74
Shrubs	1405.19	4.37	3067.56	8.22
Fields	9163.56	28.47	12961.25	34.75
Crops	0	0	9326.50	25.01
Dense urban development	52.88	0.16	1957.50	5.25
Low urban development	70.13	0.22	811.13	2.17
Parkings, quarries	12.06	0.04	97.31	0.26
Rocks	108.25	0.34	0	0
Water	1029.63	3.20	3156.19	8.46

Land use/land cover was extracted for each atlas cell from a 1988 classified Landsat TM image with a 25-m pixel resolution. The supervised classification was labelled by intensive ground-truthing of a topographically diverse 15 by 15 km^2 area and by comparison with existing, fine-scale land cover maps (Vuillod 1994), and resulted in 12 land cover classes (Table 1). I simplified this classification scheme by grouping all non-forest classes, which brought down the number of classes to only five: oak-hornbeam-beech (Decid), pure beech (Beech), beech-fir-spruce (Conif), conifer plantations (Planted), and not forested (Open). I computed edge density (Edgeden) between forest and non-forest patches for each cell after grouping the first four classes into one, and vectorizing the resulting file (see methods in Chapter 2). I obtained mean elevation (Meanelev) for each cell by averaging elevation values from a 50 m digital elevation model purchased from the French Institut Geographique National.

Modeling procedures

I used multiple logistic regression (LR) to create models to classify the presence and absence of the green woodpecker. To prevent multicollinearity, which could artificially raise classification accuracy, I computed Pearson's product-moment correlation coefficient (r) between all pairs of variables and eliminated one variable from pairs with r greater than 0.7 (Green 1979). The decision about which variable to eliminate was based on the results of univariate LR (log likelihood and Wald statistics; SPSS 1990). Parsimonious models were developed from the remaining pool of variables by using both forward and backward stepwise selection procedures. Addition of variables in the forward procedure was based on the Wald statistic, using a P-of-entry of 0.05. Removal of variables in the backward procedure was also based on the Wald statistic, but a P-of-removal of 0.1 was used. Although stepwise procedures have

been criticized (Johnson 1981, James and McCulloch 1990), I chose to use them because the pool of variables from which the models were built was already small, and because these procedures provided an objective, repeatable approach to model building. When the two procedures resulted in different models, I retained the best of the two models based on log likelihood, Wald statistics of the predictor variables, and improvement of the model over chance classification as estimated from Cohen's Kappa (*K*) (Titus *et al.* 1984; Chapter 1). Because the output of LR is probabilistic, allocation of cases to predicted groups (presence or absence) required that a cut-off be defined; I retained the mid-point between the mean probabilities for the presence and absence cells (Fielding and Haworth 1995). Even though this rule may not maximize Kappa, it was adopted because of its objectivity and consistency. This basic modeling approach was repeated for each of the following investigations:

Influence of geographic location

The ability of models developed for the Orbe Valley and for Geneva Canton to correctly predict green woodpecker distribution in a different geographic area was tested by applying each model to the other site. These two models will be referred to as "full" by contrast to the subset models described below.

Influence of number of cells

The ratio of green woodpecker presence-to-absence cells differed between the two study sites (49/224 = 0.22 in the Orbe Valley, 203/103 = 1.97 in Geneva Canton). To correct this difference, I randomly selected 33 presence and 44 absence cells in each area, to obtain a ratio of 0.75 for both sites. This ratio was preferred over a 1.0 ratio because a higher number of absence cells can be an advantage because absences are

likely to exhibit more variability (Capen *et al.* 1986, Pereira and Itami 1991). I selected 33 presence cells because I did not want to develop models using more than two-thirds of the presence cells in the Orbe Valley (49), and using a smaller number of presence cells would have resulted in a small sample size, which is more likely to cause model instability (Capen *et al.* 1986). Models developed from the subsets of cells were used to predict green woodpecker presence in the unused cells within each study site, and in all the cells of the other site.

Influence of MMU

I used a merge program to resample the unsupervised classification to a 1-ha MMU using a rule-based algorithm (Ma 1995). Previous manipulations showed that increasing the MMU to 2 and 4 ha resulted in little additional changes (Chapter 2). I developed models using the new image and tested each model on the other site.

Influence of atlas cell size

I created new, scaled-up distribution maps by grouping four 1-km atlas cells. If at least one of the four cells was coded as presence, the new, 400-ha cell was coded as presence. Because this coding depended on which cells were aggregated, four maps were created to cover all the possible allocations of 100-ha cells (Fig. 2a and 2b). Aggregates that had only 2 or 3 cells (along the edge of the study areas) were dropped from analysis. Unfortunately, the high proportion of presence cells in Geneva Canton resulted in almost no 400-ha cells being coded as absence (Fig. 2b), so models could only be built for the Orbe Valley. For this site, the presence/absence ratio varied from 0.84 to 1.3. I developed models from each distribution map and applied them to the four scaled-up distribution maps of Geneva Canton.

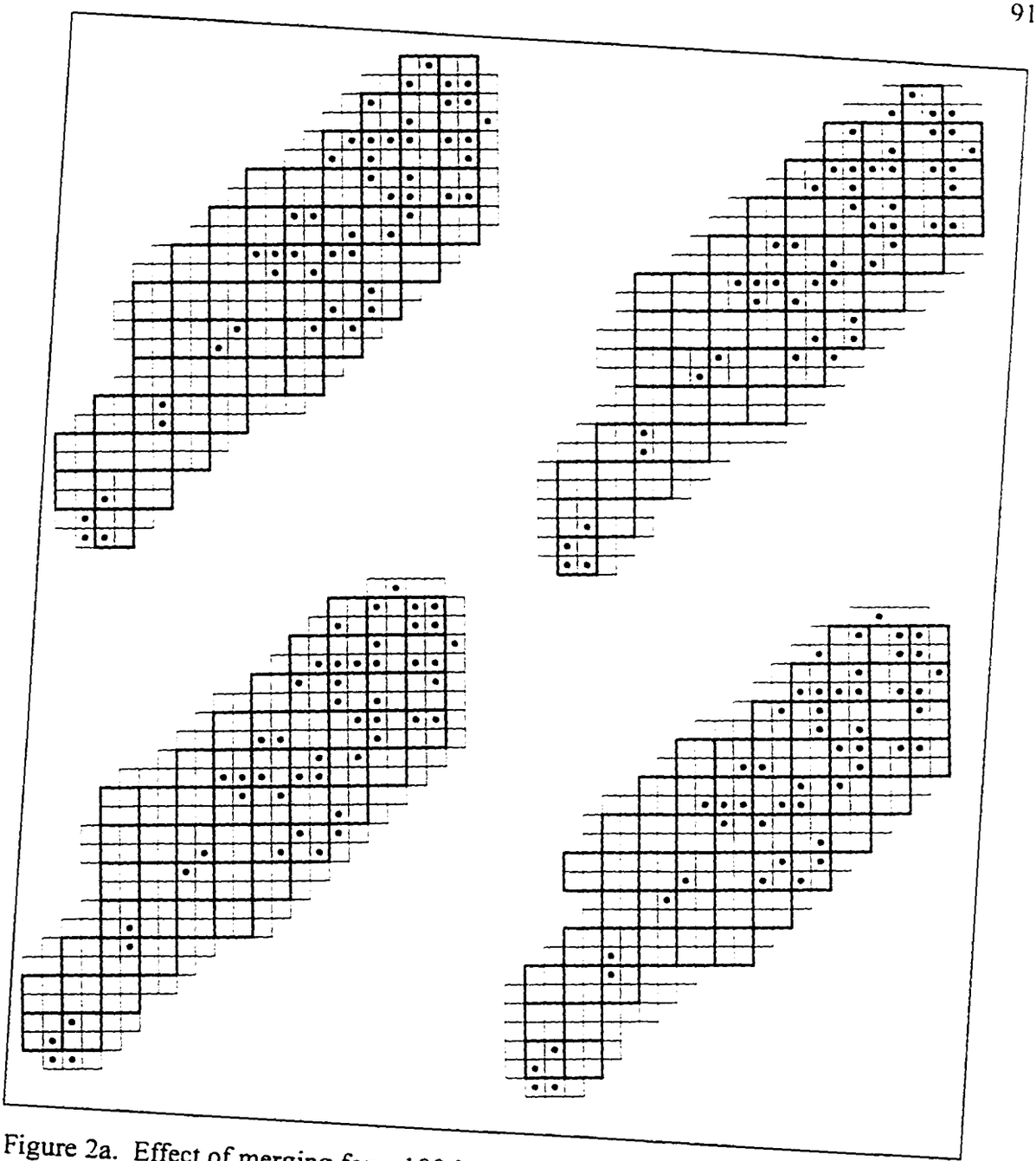


Figure 2a. Effect of merging four, 100-ha breeding bird atlas cells into one, 400-ha cell, on the distribution of presences and absences of the green woodpecker in the Orbe Valley, Switzerland. Black dots: presences in 100-ha cells; shaded squares: presences in 400-ha cells.

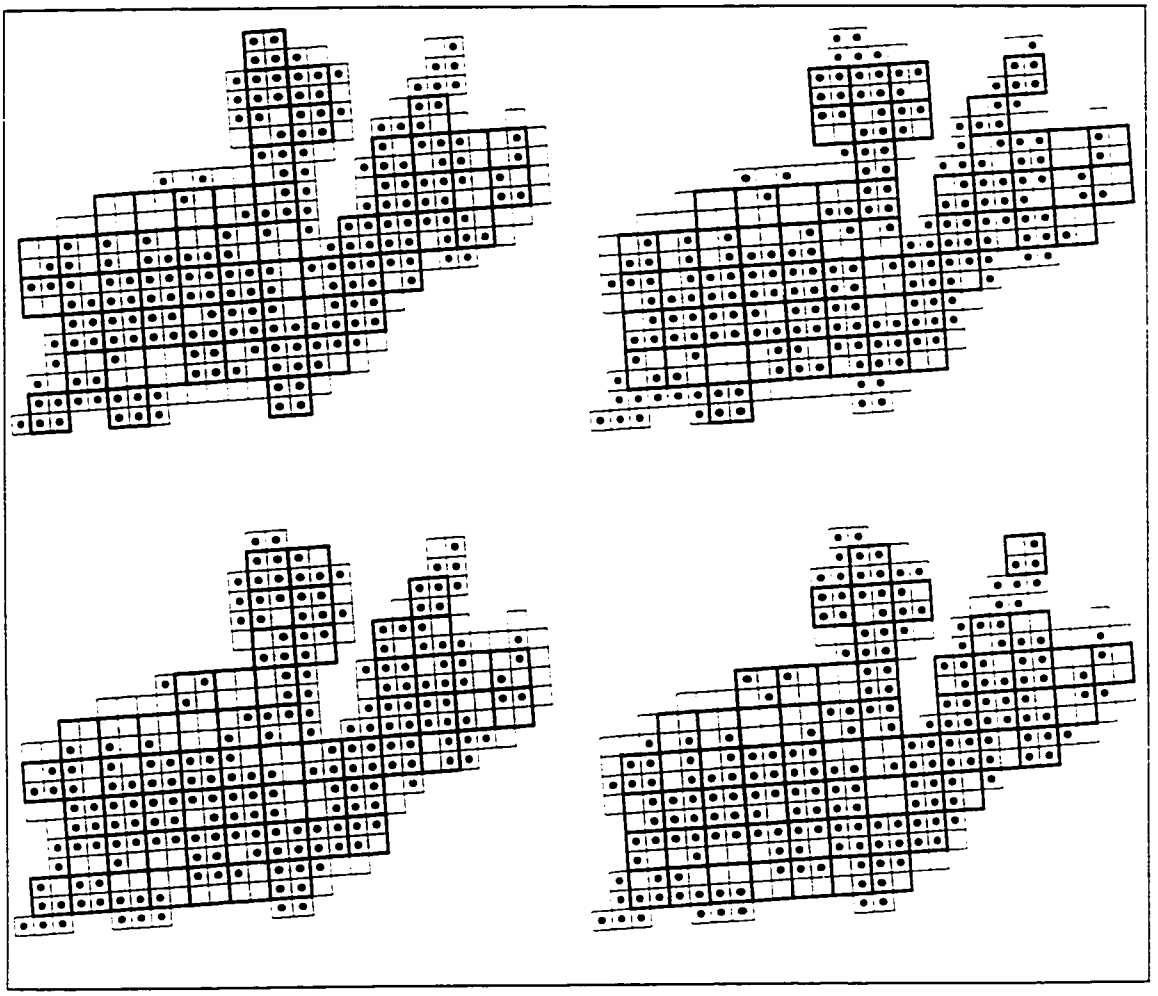


Figure 2b. Effect of merging four, 100-ha breeding bird atlas cells into one, 400-ha cell, on the distribution of presences and absences of the green woodpecker in Geneva Canton, Switzerland. Black dots: presences in 100-ha cells; shaded squares: presences in 400-ha cells.

Combined influence of MMU and atlas cell size

I repeated the procedure described above, using the 1-ha MMU image.

RESULTS

Selection of variables

For all treatments, the spatial distribution of landscape patches drove the selection of habitat variables. In Geneva Canton, edge density values were low for cells located in the agricultural matrix (< 20 m/ha), and reached their highest values (> 200 m/ha) in those cells that overlapped one of the scattered forest patches. This resulted in high correlation coefficients between Edgeden and Beech, Decid, Conif, Planted, and Open (Table 2a). Correlations between land cover classes and Edgeden remained high when the 1-ha MMU image was used, except for Conif and Planted ($r = 0.35$ and 0.31, respectively; Table 2a). Using the 0.7 criterion and univariate LR results (Table 3), I retained Meanelev, Edgeden, and Conif to build the 0.0625-ha MMU models, and Meanelev, Edgeden, Conif, and Planted to build the 1-ha MMU models. Correlation coefficients for the 10 subsets of cells were similar to those presented in Table 2a (0.0625-ha MMU), so Meanelev, Edgeden and Conif were selected to enter the stepwise procedures.

In the Orbe Valley, the contiguity of the mixed forest (especially on the northwestern slope of the valley; see Fig. 2, Chapter 2) resulted in extremely strong negative correlations between Conif and Open, at both MMUs, for both atlas cell sizes, and for all 10 subsets of cells ($|r| > 0.97$; Tables 2b, 6a, 6b). Wald statistic values were slightly higher (and the corresponding P-values slightly lower) for Conif than for Open in univariate LR analyses (Tables 3 and 7a), except with 400-ha cells at 1-ha MMU

Table 2a. Spearman rank correlations among seven variables in Geneva Canton, Switzerland, at two different minimum mapping units (MMU).

MMU (ha)		Edgeden	Meanelev	Decid	Beech	Conif	Planted
0.0625	Meanelev	-0.212					
1		-0.023					
0.0625	Decid	0.769	-0.054				
1		0.947	-0.058				
0.0625	Beech	0.911	-0.177	0.786			
1		0.702	-0.022	0.591			
0.0625	Conif	0.685	-0.257	0.577	0.734		
1		0.346	-0.031	0.279	0.415		
0.0625	Planted	0.711	-0.241	0.636	0.676	0.666	
1		0.313	0.041	0.197	0.304	0.362	
0.0625	Open	-0.881	0.111	-0.943	-0.928	-0.930	-0.731
1		-0.986	-0.011	-0.964	-0.695	-0.363	-0.319

Table 2b. Spearman rank correlations among six variables in the Orbe Valley, Switzerland, at two different minimum mapping units (MMU).

MMU (ha)		Edgeden	Meanelev	Beech	Conif	Planted
0.0625	Meanelev	0.231				
1		0.167				
0.0625	Beech	0.311	-0.030			
1		0.107	-0.210			
0.0625	Conif	-0.431	0.431	-0.288		
1		-0.421	0.446	-0.261		
0.0625	Planted	0.018	-0.011	0.173	0.324	
1		0.030	-0.040	0.018	0.104	
0.0625	Open	0.447	-0.383	0.107	-0.971	-0.375
1		0.449	-0.409	0.066	-0.970	-0.016

Table 3. Univariate logistic regression results (-2 Log Likelyhood, regression coefficient, Wald statistic and corresponding *P*-value) for seven variables at two different minimum mapping units (MMU), for distribution datasets of the green woodpecker in the Orbe Valley and in Geneva Canton, Switzerland.

Variable	Model (MMU)	-2LL	Coefficient	Wald	P
Meanelev	Orbe (0.0625)	239.73	-0.0052	15.20	0.0001
	Orbe (1)	239.73	-0.0052	15.20	0.0001
	Geneva (0.0625)	373.15	-0.0268	15.27	0.0001
	Geneva (1)	373.15	-0.0268	15.27	0.0001
Edgeden	Orbe (0.0625)	256.88	0.0009	0.07	0.7891
	Orbe (1)	250.96	0.0122	5.87	0.0154
	Geneva (0.0625)	354.96	0.0173	29.06	0.0000
	Geneva (1)	389.35	0.0068	1.50	0.2207
Decid	Geneva (0.0625)	387.80	0.0280	2.79	0.0951
	Geneva (1)	390.67	0.0061	0.25	0.6194
	Orbe	N/A	N/A	N/A	N/A
Beech	Orbe (0.0625)	256.36	0.0179	0.62	0.4323
	Orbe (1)	256.93	0.0039	0.03	0.8610
	Geneva (0.0625)	383.54	0.0611	5.65	0.0175
	Geneva (1)	390.79	0.0069	0.12	0.7254
Conif	Orbe (0.0625)	214.26	-0.0240	14.21	0.0002
	Orbe (1)	242.20	-0.0202	13.85	0.0002
	Geneva (0.0625)	375.18	0.3896	9.19	0.0024
	Geneva (1)	387.08	0.2794	1.10	0.2954
Planted	Orbe (0.0625)	251.48	0.2114	5.53	0.0187
	Orbe (1)	251.93	1.1541	2.08	0.1496
	Geneva (0.0625)	356.41	1.4160	23.04	0.0000
	Geneva (1)	389.48	0.2442	1.08	0.2989
Open	Orbe (0.0625)	244.08	0.0210	12.36	0.0004
	Orbe (1)	242.46	0.0202	13.82	0.0002
	Geneva (0.0625)	382.12	-0.0271	7.36	0.0067
	Geneva (1)	390.27	-0.0069	0.63	0.4289

(Table 7b). For consistency, models for the Orbe Valley were generated for all investigations from a reduced pool of predictor variables composed of Meanelev, Edgeden, Beech, Conif, and Planted.

Classification results

Geographic location, number of cells, MMU and atlas cell size all influenced the composition of the regression equation and the classification accuracy of the models predicting green woodpecker presence and absence. Meanelev and Edgeden entered the full model for Geneva Canton at 0.0625-ha MMU (Table 4), and four of the ten subset models; Edgeden was the sole variable retained by the other six models (Table 5a). Improvement over chance classification (K) at 0.0625-ha MMU varied from 17.50% to 47.80%; absences were usually better predicted than presences (Tables 4 and 5a). In the Orbe Valley, using a subset of cells resulted in more parsimonious models for seven of the ten subsets (Table 5b), and this improved classification success by increasing the number of presences correctly classified (from 61.22% for the full model, to an average of 70.98% for the subset models).

Increasing the MMU to 1 ha had a strong influence on edge density values, and reversed the significance of Edgeden in univariate LR (from non-significant to significant in the Orbe Valley, and from significant to non- significant in Geneva Canton, even using a P-value as high as 0.2; Table 3). In Geneva Canton, the Meanelev-only model resulted in fewer misclassifications among the presences (83.7% correctly classified; Table 4), but fewer absences were correctly classified than with the 0.0625 ha MMU model (36.9% vs 68.9%; Table 4), so overall model performance as measured by Cohen's Kappa was lower at 1-ha MMU (0.223 vs 0.347; Table 4). In the Orbe Valley model, Conif was replaced by Edgeden, but this had little effect on

Table 4. Regression equations, corresponding statistics (standard error of the estimate *SEE*, coefficient of determination R^2), and classification results (percent presence and absence correctly classified, Cohen's Kappa and corresponding *Z*-value) for logistic models developed at two different minimum mapping units (MMU) from green woodpecker distribution data in the Orbe Valley and in Geneva Canton, Switzerland; and classification success of the models when tested with the other sites' data. Neg.: Negative value. * = $P \leq 0.005$.

					% correctly classified			
MMU (ha)	Model/test	Regression equation	*SEE*	R^2	Presence	Absence	Kappa	Z-value
0.0625	Orbe Valley	2.0957 - 0.0029*Meanelev - 0.0284*Conif + 0.3865*Planted	0.363	0.230	61.22	69.20	0.218	3.008*
	Geneva Canton	8.7277 - 0.0218*Meanelev + 0.0158*Edgeden	0.436	0.626	68.47	68.93	0.347	5.849*
	Orbe on Geneva	N/A	-	-	56.16	82.52	0.331	6.050*
	Geneva on Orbe	N/A	-	-	0	100	Neg.	N/A
1	Orbe Valley	4.7992 - 0.0061*Meanelev + 0.02*Edgeden + 1.0318*Planted	0.361	0.238	61.22	72.32	0.250	3.339*
	Geneva Canton	12.0752 - 0.0291*Meanelev + 0.292*Conif	0.475	0.590	83.74	36.89	0.223	3.275*
	Orbe on Geneva	N/A	-	-	42.36	66.99	0.078	1.465
	Geneva on Orbe	N/A	-	-	18.37	66.52	Neg.	N/A

Table 5a. Regression equations, regression statistics (standard error of the estimate *SEE*, coefficient of determination R^2), and classification results (percent presence and absence correctly classified, Cohen's Kappa and corresponding *Z*-value) for ten logistic regression models developed from green woodpecker distribution data in Geneva Canton, Switzerland; and classification success of the models (Cohen's Kappa and corresponding *Z*-value) when applied to data not used for model development (*Remaining data*). * = $P \leq 0.05$; ** = $P \leq 0.005$.

			Classification results				Remaining data	
Regression equation	*SEE*	R^2	P (%)	A (%)	Kappa	Z-value	Kappa	Z-value
-1.999 + 0.0251*Edgeden	0.451	0.514	60.61	72.73	0.335	2.866**	0.195	2.866**
-0.7754 + 0.0066*Edgeden	0.492	0.436	51.52	65.91	0.175	1.498	0.249	3.680**
-1.6358 + 0.0179*Edgeden	0.456	0.503	63.64	75.00	0.389	3.322**	0.179	2.682**
-0.8555 + 0.0078*Edgeden	0.490	0.424	57.58	63.64	0.211	1.824*	0.303	4.592**
-1.1522 + 0.0113*Edgeden	0.476	0.457	54.55	70.45	0.252	2.150*	0.231	3.415**
-2.1458 + 0.0241*Edgeden	0.440	0.536	63.64	79.55	0.437	3.714**	0.179	2.682**
10.6435 + 0.0214*Edgeden - 0.0293*Meanelev	0.429	0.564	69.70	75.00	0.445	3.843**	0.220	3.278**
7.8998 + 0.0189*Edgeden - 0.0227*Meanelev	0.449	0.524	57.58	77.27	0.354	2.997**	0.249	3.680**
12.4418 + 0.0232*Edgeden - 0.0344*Meanelev	0.426	0.571	75.76	72.73	0.478	4.168**	0.234	3.595**
11.2001 + 0.0135*Edgeden - 0.0296*Meanelev	0.453	0.515	63.64	77.27	0.412	3.518**	0.249	3.680**

Table 5b. Regression equations, regression statistics (standard error of the estimate *SEE*, coefficient of determination R^2), and classification results (percent presence and absence correctly classified, Cohen's Kappa and corresponding *Z*-value) for ten logistic regression models developed from green woodpecker distribution data in the Orbe Valley, Switzerland; and predictive accuracy of the models (Cohen's Kappa and corresponding *Z*-value) when applied to data not used for model development (*Remaining data*) and data from Geneva Canton (*Geneva Canton*). Neg.: Negative value. * = $P \leq 0.05$; ** = $P \leq 0.005$.

			Classification results				Remaining data		Geneva Canton	
Regression equation	*SEE*	R^2	P (%)	A (%)	Kappa	*Z*-value	Kappa	*Z*-value	Kappa	Z-value
0.0682 - 0.0712*Conif + 1.1089*Planted	0.392	0.637	76.47	80.00	0.508	5.064**	0.113	1.474	0.327	5.899**
0.3356 - 0.0368*Conif + 0.3171*Planted	0.473	0.472	69.70	61.36	0.303	2.670**	0.105	1.267	0.289	5.362**
0.3431 - 0.0438*Conif + 0.5003*Planted	0.441	0.541	75.76	68.18	0.430	3.777**	0.113	1.430	0.313	5.670**
-0.5885 - 0.0427*Conif + 0.7912*Planted	0.438	0.547	75.76	70.45	0.454	3.972**	0.062	0.739	0.306	5.668**
-0.2655 - 0.394*Conif + 0.5908*Planted	0.462	0.496	66.67	63.64	0.297	2.605**	0.135	1.607	0.308	5.668**
0.341 - 0.0357*Conif + 0.3492*Planted	0.458	0.504	66.67	59.09	0.251	2.214*	0.142	1.768	0.291	5.363**
0.4805 - 0.0403*Conif + 0.3323*Planted	0.457	0.506	69.70	63.64	0.326	2.865**	0.085	1.094	0.289	5.362**
6.4039 - 0.0318*Conif + 0.5996*Planted - 0.0058*Meanelev	0.435	0.559	69.70	68.18	0.373	3.256**	0.084	1.003	0.396	6.828**
6.5094 - 0.0284*Conif + 0.4016*Planted - 0.0054*Meanelev	0.427	0.575	69.70	75.00	0.445	3.843**	0.084	1.003	0.176	3.611**
1.2205 - 0.044*Conif + 0.5563*Planted - 0.0117*Edgeden	0.452	0.523	69.70	68.18	0.373	3.256**	0.111	1.318	Neg.	N/A

classification results ($K = 0.218$ at 0.0625-ha MMU vs $K = 0.250$ at 1-ha MMU; Table 4).

Spearman rank correlations, univariate LR results, and regression equations for the four models created in the Orbe Valley from the 400-ha atlas cells were similar within MMUs (Tables 6a, 6b, 7a and 7b), but the composition of the equation changed with the MMU (Conif and Planted at 0.0625 ha, Meanelev and Edgeden at 1 ha; Table 8). The significance of Edgeden increased at 1-ha MMU, whereas that of Planted decreased (Wald statistic; Tables 7a and 7b). At both MMUs, classification results were higher than when models were built from 100 ha-cells, with K-values averaging 0.464 at 0.0625-ha MMU and 0.486 at 1-ha MMU (Table 8). This increase was caused by a better prediction of green woodpecker presences for the 400-ha cell models compared to the 100-ha cell ones (from 61% to an average of 79% at 0.0625 ha MMU, and from 61% to an average of 75% at 1 ha MMU; Table 8). Classification rates of absences remained fairly constant: 69% with 100-ha cells, average of 67% with 400-ha cells at 0.0625-ha MMU; 72% with 100-ha cells, average of 74% with 400-ha cells at 1-ha MMU (Table 8).

Model universality

The ability of the models to correctly predict at least some of the other site's woodpecker presences was observed in one direction only: the full model developed from the Orbe Valley data using 100-ha cells and a MMU of 0.0625 ha correctly classified 56.16% of presences and 82.52% of absences in Geneva Canton, whereas the Geneva Canton model misclassified all the presence as absences in the Orbe Valley (Table 4). Similar results were obtained from the subset models. Improvement over chance classification was not statistically significant ($P > 0.05$) when the ten subset

Table 6a. Spearman rank correlations among six variables for four datasets in the Orbe Valley, Switzerland. Land cover variables were extracted from a classified landsat image with a 0.0625-ha minimum mapping unit.

		Edgeden	Meanelev	Beech	Conif	Planted
Meanelev	1 (N = 59)	0.355				
	2 (N = 53)	0.263				
	3 (N = 54)	0.224				
	4 (N = 57)	0.325				
Beech	1 (N = 59)	0.172	-0.095			
	2 (N = 53)	0.207	-0.016			
	3 (N = 54)	0.144	-0.140			
	4 (N = 57)	0.203	-0.057			
Conif	1 (N = 59)	-0.404	0.478	-0.181		
	2 (N = 53)	-0.350	0.582	-0.142		
	3 (N = 54)	-0.352	0.660	-0.303		
	4 (N = 57)	-0.434	0.469	-0.212		
Planted	1 (N = 59)	0.0003	-0.094	0.190	0.159	
	2 (N = 53)	0.153	0.090	0.249	0.274	
	3 (N = 54)	0.140	0.153	0.023	0.282	
	4 (N = 57)	0.174	0.057	0.243	0.262	
Open	1 (N = 59)	0.425	-0.428	0.027	-0.980	-0.205
	2 (N = 53)	0.387	-0.551	-0.033	-0.971	-0.286
	3 (N = 54)	0.380	-0.626	0.156	-0.980	-0.299
	4 (N = 57)	0.423	-0.434	0.042	-0.972	-0.356

Table 6b. Spearman rank correlations among six variables for four datasets in the Orbe Valley, Switzerland. Land cover variables were extracted from a classified landsat image with a 1-ha minimum mapping unit.

		Edgeden	Meanelev	Beech	Conif	Planted
Meanelev	1 (N = 59)	0.224				
	2 (N = 53)	0.145				
	3 (N = 54)	0.143				
	4 (N = 57)	0.248				
Beech	1 (N = 59)	-0.114	-0.375			
	2 (N = 53)	0.011	-0.261			
	3 (N = 54)	-0.020	-0.333			
	4 (N = 57)	0.078	-0.303			
Conif	1 (N = 59)	-0.387	0.507	-0.213		
	2 (N = 53)	-0.372	0.623	-0.127		
	3 (N = 54)	-0.345	0.677	-0.360		
	4 (N = 57)	-0.435	0.511	-0.206		
Planted	1 (N = 59)	0.043	-0.104	-0.201	-0.131	
	2 (N = 53)	0.149	-0.125	-0.016	-0.174	
	3 (N = 54)	0.088	0.087	-0.039	-0.048	
	4 (N = 57)	0.127	-0.033	0.007	-0.074	
Open	1 (N = 59)	0.420	-0.456	0.046	-0.980	0.167
	2 (N = 53)	0.405	-0.581	-0.069	-0.970	0.210
	3 (N = 54)	0.388	-0.650	0.224	-0.983	0.042
	4 (N = 57)	0.433	-0.461	0.029	-0.976	0.057

Table 7a. Univariate logistic regression results (-2 Log Likelyhood, regression coefficient, Wald statistics and corresponding *P*-value) for six variables, for four distribution datasets of the green woodpecker in the Orbe Valley, Switzerland. Land cover variables were extracted from a classified landsat image with a 0.0625-ha minimum mapping unit.

Variable	Model	-2LL	Coefficient	Wald	P
Meanelev	1	74.13	- 0.0060	6.27	0.0123
	2	64.82	-0.0071	6.69	0.0097
	3	67.66	-0.0064	5.80	0.0160
	4	73.62	-0.0051	4.51	0.0337
Edgeden	1	79.65	0.0088	1.66	0.1979
	2	70.80	0.0108	1.69	0.1938
	3	73.00	0.0082	1.16	0.2816
	4	77.31	0.0082	1.25	0.2641
Beech	1	80.93	0.438	0.44	0.5093
	2	72.51	0.0111	0.04	0.8442
	3	74.18	0.0061	0.01	0.9250
	4	78.42	0.0254	0.16	0.6906
Conif	1	65.08	-0.055	11.72	0.0006
	2	59.66	-0.052	9.87	0.0017
	3	63.20	-0.0434	8.95	0.0028
	4	72.26	-0.0321	5.56	0.0184
Planted	1	77.99	0.5155	3.16	0.0753
	2	69.75	0.4781	2.54	0.1109
	3	70.23	0.5339	3.47	0.0626
	4	74.42	0.5395	3.52	0.0608
Open	1	66.91	0.0504	10.93	0.0009
	2	60.28	0.0517	9.24	0.0024
	3	63.08	0.0459	8.84	0.0029
	4	73.09	0.0299	4.86	0.0276

Table 7b. Univariate logistic regression results (-2 Log Likelyhood, regression coefficient, Wald statistics and corresponding *P*-value) for six variables, for four distribution datasets of the green woodpecker in the Orbe Valley, Switzerland. Land cover variables were extracted from a classified landsat image with a 1-ha minimum mapping unit.

Variable	Model	-2LL	Coefficient	Wald	P
Meanelev	1	74.13	- 0.0060	6.27	0.0123
	2	64.82	-0.0070	6.69	0.0097
	3	67.66	-0.0064	5.80	0.0160
	4	73.62	-0.0051	4.51	0.0337
Edgeden	1	74.63	0.0291	5.78	0.0162
	2	64.86	0.0353	6.57	0.1040
	3	67.30	0.032	5.83	0.0157
	4	75.01	0.0205	3.33	0.0068
Beech	1	81.36	0.0049	0.0066	0.9354
	2	72.47	-0.0147	0.0761	0.7826
	3	73.96	-0.0276	0.2272	0.6336
	4	78.57	-0.0065	0.0119	0.9133
Conif	1	66.05	-0.0468	11.39	0.0007
	2	60.10	-0.0461	9.56	0.0020
	3	64.09	-0.0374	8.33	0.0039
	4	72.59	-0.0283	5.26	0.0218
Planted	1	79.37	3.0275	0.82	0.3647
	2	725.07	1.046	0.37	0.5433
	3	70.53	24.75	0.09	0.7676
	4	76.73	3.1314	0.83	0.3635
Open	1	66.01	0.0470	11.58	0.0007
	2	59.15	0.0498	9.85	0.0017
	3	62.34	0.0437	9.30	0.0023
	4	72.40	0.0291	5.40	0.0201

Table 8. Regression equations, corresponding statistics (standard error of the estimate *SEE*, coefficient of determination R^2), and classification results (percent presence and absence correctly classified, Cohen's Kappa and corresponding *Z*-value) for eight logistic models developed at two different minimum mapping units (MMU) from green woodpecker distribution data in the Orbe Valley, Switzerland. All the *Z*-values have corresponding *P*-values smaller than 0.005.

					% correctly classified			
MMU	Model	Regression equation	*SEE*	R^2	Presence	Absence	Kappa	*Z*-value
0.0625	1	0.4966 - 0.0736*Conif + 1.0409*Planted	0.413	0.628	77.78	68.75	0.460	3.548
	2	0.4969 - 0.0715*Conif + 1.1407*Planted	0.416	0.696	83.33	65.22	0.493	3.475
	3	0.39 - 0.0678*Conif + 1.1775*Planted	0.411	0.697	80.00	70.83	0.510	3.689
	4	-0.0127 - 0.0439*Conif + 0.8664*Planted	0.449	0.629	77.42	61.54	0.393	2.919
1	1	8.9399 - 0.0094*Meanelev + 0.0497*Edgeden	0.424	0.608	77.78	75.00	0.525	4.025
	2	10.1838 - 0.0101*Meanelev + 0.0516*Edgeden	0.415	0.697	80.00	69.57	0.498	3.546
	3	9.6629 - 0.0095*Meanelev + 0.0466*Edgeden	0.422	0.680	70.00	79.17	0.484	3.568
	4	7.7342 - 0.0074*Meanelev + 0.0333*Edgeden	0.454	0.620	70.97	73.08	0.438	3.299

Table 9. Cohen's Kappa and corresponding *Z*-value obtained from applying four logistic models developed from green woodpecker distribution data in the Orbe Valley, Switzerland (column), to four datasets in Geneva Canton, Switzerland (row). Models were developed from a Landsat classified image with a 0.0625-ha minimum mapping unit. All the *Z*-values have corresponding *P*-values greater than 0.1.

		1	2	3	4
1	Kappa	0.206	0.036	0.118	0.171
	Z-value	1.128	0.274	0.966	0.837
2	Kappa	0.237	0.036	0.268	0.139
	Z-value	1.212	0.274	1.490	0.751
3	Kappa	0.237	0.046	0.133	0.139
	Z-value	1.212	0.327	1.033	0.751
4	Kappa	0.181	0.036	0.133	0.129
	Z-value	1.054	0.274	1.033	0.724

models for the Orbe Valley were applied to the remaining cases (Table 5b), but nine of the ten models improved chance classification by 17.6% to 39.6% when applied to Geneva Canton; only the model including Edgeden resulted in a negative K-value (Table 5b). Conversely, the subset models for Geneva Canton performed better when applied to the remaining cases (K ranging from 17.9% to 30.3%, $P \leq 0.005$, Table 5a) than they did with data from the Orbe Valley, where none of the ten models provided improvement over chance classification ($K = 2.0\%$ for one model, and $K < 0$ for the other nine models).

Increasing the MMU of the Landsat image did not improve predictive generality; the Geneva Canton model still misclassified all the presences in the Orbe Valley, and the Orbe Valley model predicted only 42.4% of the presences and 67.0% of the absences in Geneva Canton, which yielded a lower improvement over chance classification than that obtained from the 0.0625-ha MMU model (Table 4).

Performance of the 400-ha cells, 0.0625-ha Orbe models varied among the four, scaled-up distribution datasets of Geneva Canton. The four Orbe models correctly classified an average of 73.7% of the presences, and 100% of the absences for the first Geneva dataset, versus 52.7% and 66.7% for the second dataset, 53.6% and 100% for the third dataset, and 70.69% and 100% for the fourth dataset. The overall lower prediction rates for the second Geneva dataset resulted in lower K values for that dataset (Table 9). The models, however, failed to provide a statistical improvement over chance classification, regardless of the dataset to which they were applied ($P > 0.1$; Table 9). Finally, the four, 1-ha MMU models of the Orbe Valley performed poorly and classified all the 400-ha cells of the Geneva Canton datasets as presences.

DISCUSSION

My aims were 1) to determine if logistic regression models developed for the green woodpecker in two Swiss areas using simple, broad-scale habitat variables exhibited model universality; and 2) to assess how changes in scale of the habitat and distribution data affected model classification results and model universality. The results suggest that model performances were variable and a function of the characteristics of the landscapes, the MMU of the habitat map, and the size of the atlas distribution grid cells.

Geographic generality

Consistent patterns of associations between a species and characteristics of its habitat, as well as the ability to use these characteristics as variables, are requisites of model universality. The simple variables used in this study proved unsuitable for predicting distribution of the green woodpecker. The models developed for the Orbe Valley and Geneva Canton comprised different variables (Table 4), whether all the presence/absence cells or only a subset of them were used for analysis. Although species-habitat associations may vary geographically (e.g. Collins 1983, Shy 1984), these differences are more likely an artifact caused by the scale of the study and the variables used. In the Orbe Valley, the presence of the green woodpecker was negatively associated with the variable Conif and, conversely, positively associated with the variable Open, and there was no significant correlation with edge (univariate LR; Table 3). The exact opposite was found in Geneva Canton: positive correlation with Conif, negative one with Open, and strong significance of the variable Edgeden ($P <$ 0.0005; Table 3). This apparent contradiction disappears when the structure and

composition of the entire landscapes are considered, instead of just the composition of individual atlas cells. Indeed, forest patch characteristics differ between the two sites. In the Orbe Valley, they tend to be large and unbroken. The green woodpecker is known to avoid closed, dense coniferous forests, favoring instead open or broken deciduous or mixed forests with grassy fringes or clearings (Short 1982, Cramp 1985, Spitznagel 1989, Hågvar *et al.* 1990, Angelstam and Mikusinski 1994); hence, the negative correlation with Conif. By contrast, in Geneva Canton, forest patches are smaller and scattered in the agricultural matrix. Although considered to be more an arboreal than a forest species (Cramp 1985), the green woodpecker still requires forest patches. Hence, the positive correlation between the species' presence and forest classes (and the correlated variable Edgeden) in Geneva Canton. Because the models did not incorporate patch configuration attributes such as patch size, fundamental differences between the two sites could not be taken into account during the modeling phase.

Scale has been defined as the interaction of grain and extent, where grain relates to the level of resolution (i.e., MMU), and extent to the largest entities that can be detected in the data (size of the study area or duration of time under consideration) (Allen and Hoekstra 1991, Turner *et al.* 1989, 1993). Using this definition, the Orbe Valley and Geneva Canton study sites were at similar scales; however, because of the presence of larger forest patches, the Orbe valley can be considered a "coarse-grained" landscape, compared to Geneva Canton (Forman and Godron 1986). This suggests that the two datasets may in fact have been at different scales, which probably contributed to the poor generality of the models. Spatial characteristics of patches may have been better predictors of the species' presence within and between sites. Spatially-explicit models that incorporate information about patch size and arrangement (Van Horne 1990) are

likely to have higher predictive capabilities than composition-based models, because landscape patterns exert a strong influence on species' distribution (Hansen and Urban 1992, Gustafson *et al.* 1994, Lescourret and Genard 1994, Farina 1997). Woodpeckers, because of their large territories, are likely to be affected by the spatial patterning of the landscape (Angelstam 1989). Unfortunately, gridded data are poorly suited to extracting configuration variables such as patch size (Chapter 1).

Influence of MMU

Increasing the MMU of the habitat map from 0.0625 ha to 1 ha affected the two landscapes differently. Because dispersed land cover types tend to be lost faster than clumped types with increasing MMU (Turner *et al.* 1989, Turner 1990), many forest patches disappeared in Geneva Canton; Meanelev remained as the only statistically significant variable in univariate LR (Table 3). Although the green woodpecker tends to select lower-elevation sites for nesting (Glue and Boswell 1994), there may not be enough topographic relief within Geneva Canton for this variable to be a strong predictor of the species' presence. In fact, a univariate LR model of elevation could not be developed at either MMUs, because the cut-off was 1. However, a model with edge density alone correctly classified 66.01% of presences, and 66.99% of absences, at 0.0625-ha MMU. Thus, in the model combining Meanelev and Edgeden (0.0625-ha MMU model; Table 4), the classification success was mostly due to Edgeden. In the 1-ha MMU model, the replacement of Edgeden by Conif resulted in lower classification success, probably because Conif was a poor predictor of the species' presence (Table 3); only a small proportion of the landscape was classified as beech-fir-spruce forests (1.59%, Table 1; dropped to 0.48% at 1-ha MMU, Chapter 2: Table 3b).

In the Orbe valley, the elimination of small groups of pixels may have clarified

patterns that were obscured at the finer MMU; the significant, positive correlation with edge density at that scale is in agreement with the description of the green woodpecker as an edge species (Short 1982, Clergeau and Chefson 1988). Replacing Conif in the equation of the 0.0625-ha MMU model with Edgeden at 1-ha MMU had little effect on classification results within the Orbe valley, but reduced the universality of the model when applied to Geneva Canton (Table 4), because Edgeden was a non-significant predictor variable in that area (Table 3).

Influence of grid cell size

The choice of the grid cell size for breeding bird atlases (and other distribution atlases) is a compromise between the level of detail sought and the manpower available to conduct censuses. A 100-ha cell size was retained for both areas, but for larger sites even this coarse a sampling may not be possible (e.g., Jovéniaux 1993). In the Orbe Valley, increasing the cell size to 400 ha almost doubled classification success, possibly by clarifying bird-habitat association patterns. Heikkinen (1988) suggested that distribution patterns of rare plant species richness in a Finnish reserve may have been more obscured at the 1-km grid scale he used for his models, than at either finer or broader scales. Unfortunately, the number of 400-ha absence cells in Geneva Canton was too small (from 2 to 4; Fig. 2b) to allow models to be developed, so it was not possible to assess whether the classification improvement observed in the Orbe Valley was site-specific, or a more general pattern.

That green woodpecker presence was predicted for almost all 400-ha cells in Geneva Canton demonstrates the influence of scale in data collection procedures: at 400 ha, all of Geneva Canton appeared suitable for the nesting green woodpecker; but at 100 ha, absence cells were more numerous. Without first-hand knowledge of the area, it is

difficult to know whether these absence cells were comprised of truly unsuitable habitat. There can be three reasons for absence cells: the cell lacks suitable habitat; the cell contains suitable habitat, but was unoccupied during the time frame of the census (this is especially likely for species exhibiting metapopulation dynamics); or, the species was present, but went undetected. Presenting distribution data as probabilities of occurrence, such as the output of LR models built from the original presence/absence data, is one way to limit the problem of "false absences" (Osborne and Tigar 1992).

The proportion of cells in which green woodpeckers were predicted to be present also increased in the Orbe Valley, where the ratio of presences over absences reversed from 0.22 at the 100-ha scale, to an average of 1.15 at the 400-ha scale. The loss of information resulting from aggregating distribution squares could have been lessened by using an index of abundance, i.e., the number of 100-ha cells in each 400-ha cell in which the species was recorded, as input to the LR procedure (Gates *et al.* 1994).

CONCLUSION

This study suffered from several limitations. The Landsat image used to obtain the land cover variables had not been ground-truthed in either of the two study sites, and its accuracy in these sites is unknown; however, the size of the atlas cells was probably large enough, in relation to the MMU, to make the analysis relatively insensitive to misclassified pixels. More problematic is the possible presence of spatial autocorrelation in the distribution and habitat data. No attempt was made at quantifying it; but its removal usually entails eliminating neighboring cells from the analysis (e.g., Peirera and Itami 1991, Gates *et al.* 1994), and small sample sizes tend to increase model instability (Capen *et al.* 1986). It is also unlikely that removal of spatial

dependency in the data would have significantly improved model universality. Fielding and Haworth (1995), working with three bird species in five sites, also obtained variable success when they applied models developed in one area to another, even though spatial autocorrelation in their study area was negligible. Finally, the nature of the distribution data (gridded format) did not allow me to test the hypothesis that patch configuration variables were better predictor of the species' presence than just patch composition.

Despite these shortcomings, the results revealed the highly variable nature of species-habitat relationships. Because of this variability, caution is advised when applying models developed in one area to another, especially if the two landscapes are different. Mladenoff and Sickley (1998) used a logistic regression model based on road abundance in the Lake States of the Midwest, to predict potential suitable habitat for the gray wolf (*Canis lupus lycaon*) in the northeastern United States. Their model, however, may not be suitable to predict wolf habitat in the Rocky Mountains, where wolves do not appear to avoid areas of high road density (Boyd-Heger 1997). Developing models along a gradient of landscapes (instead of using only two extremes, as was done in my study) may shed more light on the predictive abilities of broad-scale models, although the results of such a study are likely to be species-specific.

REFERENCES

Allen, F. T. H., and T. W. Hoekstra. 1991. Role of heterogeneity in scaling in ecological systems under analysis. Pages 47-68 *in* J. Kolasa and S. T.A. Pickett, eds. Ecological Heterogeneity. Ecol. Studies 86. Springer-Verlag, New York, NY.

Angelstam, P. 1990. Factors determining the composition and persistence of local woodpecker assemblages in taiga forest in Sweden - a case for landscape ecological studies. Pages 147-164 in A. Carlson and G. Aulén, eds. Conservation and Management of Woodpecker Populations. Report 17, Swedish Univ. of Agric. Sci., Uppsala.

Angelstam, P, and G. Mikusinski. 1994. Woodpecker assemblages in natural and managed boreal and hemiboreal forest - a review. Ann. Zool. Fennici **31**, 157-172.

Boyd-Heger, D. K. 1997. Dispersal, genetic relationships, and landscape use by colonizing wolves in the Central Rocky Mountains. Ph.D. Dissertation, Univ. of Montana, Missoula.

Capen, D. E., J. W. Fenwick, D. B. Inkley, and A. C. Boynton. 1986. Multivariate models of songbird habitat in New England forests. Pages 171-175 *in* J. Verner, M. L. Morrison, and C. J. Ralph, eds. Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates. Univ. of Wisconsin Press, Madison.

Clergeau, P., and P. Chefson. 1988. Les Pics. Série "Comment vivent-ils" Vol. 20. Atlas Visuels Payot Lausanne. Payot Lausanne, Lausanne.

Collins, S. L. 1983. Geographic variation in habitat structure of the black-throated green warbler (*Dendroica virens*). Auk **100**, 382-389.

Converse, K. A., and B. J. Morzuch. 1981. A descriptive model of snowshoe hare habitat. Pages 232-241 *in* D. E. Capen, ed. The use of multivariate statistics in studies of wildlife habitat. USDA Fro. Serv. Gen. Tech. Rep. RM-87. Rocky Mtn. Forest and Range Exp. Sta., Fort Collins, CO.

Cramp, S. 1985. The Birds of Western Palaearctic, Vol. IV. Oxford Univ. Press, Oxford.

ESRI. 1995. Arc/Info 7.0.3. Redlands, CA.

Farina, A. 1997. Landscape structure and breeding bird distribution in a sub-Mediterranean ecosystem. Landscape Ecol. **12**, 365-378.

Forman, R. T. T., and M. Godron. 1986. Landscape Ecology. Wiley, New York, NY.

Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conserv. Biol. **9**, 1466-1481.

Gates, S., D. W. Gibbons, P. C. Lack, and R. J. Fuller. 1993. Declining farmland bird species: modelling geographical patterns of abundance in Britain. Pages 153-177 *in* P. J. Edwards, R. M. May, and N. R. Webb, eds. Large-Scale Ecology and Conservation Biology. Blackwell Sci. Publ., London.

Géroudet, P., C. Guex, M. Maire, and others. 1983. les Oiseaux Nicheurs du Canton de Genève. Muséum de Genève, Suisse.

Glayre, D., and D. Magnenat. 1984. Oiseaux Nicheurs de la Haute Vallée de l'Orbe. Nos Oiseaux 398. Fascicule spécial du volume 37.

Glue, D. E., and T. Boswell. 1994. Comparative nesting ecology of the three British

breeding woodpeckers. British Birds **87**, 253-269.

Green, R. 1979. Sampling design and statistical methods for environmental biologists. Wiley, New York, NY.

Gustafson, E. J., G. R. Parker, and S. E. Backs. 1994. Evaluating the spatial pattern of wildlife habitat: a case study. Am. Midl. Nat. **131**, 24-33.

Hågvar, S., G. Hågvar, and E. Monness. 1990. Nest site selection in Norwegian woodpeckers. Holarctic Ecol. **13**, 156-165.

Hansen, A. J., and D. L. Urban. 1992. Avian response to landscape pattern: the role of species' life history. Landscape Ecol. **7**, 163-180.

Heikkinen, R. K. 1998. Can richness patterns of rarities be predicted from mesoscale atlas data? A case study of vascular plants in the Kevo reserve. Biol. Conserv. **83**:133-143.

James, F. C., and C. E. McCulloch. 1990. Multivariate analysis: panacea or Pandora's box? Annu. Rev. Ecol. Sys. **21**, 129-166.

Jovéniaux, A. 1993. Atlas des Oiseaux Nicheurs du Jura. G.O.J. ed., Lons-le-Saunier.

Johnson, D.H. 1981. The use and misuse of statistics in wildlife habitat studies. Pages 11-19 *in* D.E. Capen, ed. The use of multivariate statistics in studies of wildlife habitat. USDA For. Serv. Gen. Tech. Rep. RM-87. Rocky Mtn. Forest and Range Exp. Sta., Fort Collins, CO.

Judson, O. P. The rise of the individual-based model in ecology. Trends in Ecol. and Evol. **9**, 9-14.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecol. **73**, 1943-1967.

Lescourret, F., and M. Genard. 1994. Habitat, landscape and bird composition in mountain forest fragments. J. Environ. Manage. **40**, 317-328.

Ma, Z. 1995. Using a rule-based merging algorithm to eliminate 'salt/pepper' and small regions of classified image. Ninth Annual Symposium on Geographic Information Systems, Vancouver, B.C., Canada, 834-837.

Mladenoff, D. J., and T. A. Sickley. 1998. Assessing potential gray wolf restoration in the northeastern United States: a spatial prediction of favorable habitat and potential population levels. J. Wildl. Manage. **62**, 1-10.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1992. Wildlife-Habitat Relationships. Univ. of Wisconsin Press, Madison.

Pereira, J. M. C., and R. M. Itami. 1991. GIS-based habitat modeling using logistic regression: a study of the Mt. Graham red squirrel. Photogrammetric Engineering and Remote Sensing **57**, 1475-1486.

PNRHJ. 1988. Le haut-Jura... De Crêts en Combes. Parc Naturel Régional du Haut-Jura, Lajoux.

Osborne, P. E., and B. J. Tigar. 1992. Interpreting bird atlas data using logistic models: an example from Lesotho, Southern Africa. J. Appl. Ecol. **29**, 55- 62.

Robertson, A., R. E. Simmons, A. M. Jarvis, and C. J. Brown. 1994. Can bird atlas data be used to estimate population size? A case study using Namibian endemics. Biol. Conserv. **71**, 87-95.

Schamberger, M., A. H. Farmer, and J. W. Terrel. 1982. Habitat suitability index models: an introduction. USDI Fish and Wildl. Serv. FWS/OBS-82/10.

Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Butterfield, S. Caicco, C. Groves, T.

Edwards, Jr., J. Ulliman, H. Anderson, F. D'Herchia, and R. G. Wright. 1993. Gap analysis: a geographic approach to protection of biodiversity. Wildl. Monogr. 123.

Short, L. L. 1982. Woodpeckers of the World. Delaware Museum of Natural History Monograph Series No. 4. Delaware Museum of Natural History, Greenville.

Shy, E. 1984. Habitat shift and geographic variation in North American tanagers (Thraupinea: *Piranga*). Oecologia **63**, 281-285.

Spitznagel, A. 1990. The influence of forest management on woodpecker density and habitat use in floodplain forests of the Upper Rhine Valley. Pages 177-146 *in* A. Carlson and G. Aulén, eds. Conservation and Management of Woodpecker Populations. Report 17, Swedish Univ. of Agric. Sci., Uppsala.

SPSS, Inc. 1990. SPSS version 4.0.4. for the Macintosh. Chicago.

Stoms, D.M. 1992. Effects of habitat map generalization in biodiversity assessments. Photogrammetric Engineering and Remote Sensing **58**, 1587-1591.

Thomas, C. D., and J. C. G. Abery. 1995. Estimating rates of butterfly decline from distribution maps: the effect of scale. Biol. Conserv. **73**, 59-65.

Titus, K., J. A. Mosher, and B. K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: ecological applications. Am. Midl. Nat. **111**, 1-7.

Turner, M. G. 1990. Spatial and temporal analysis of landscape patterns. Landscape Ecol. **4**, 21-30.

Turner, M. G., R. V. O'Neill, R. H. Gardner, and B. T. Milne. 1989. Effects of changing spatial scale on the analysis of landscape patterns. Landscape Ecol. **3**, 153-162.

Turner, M. G., W. H. Romme, R. H. Gardner, R. V. O'Neill, and T. H. Kratz. 1993. A revised concept of landscape equilibrium: disturbance and stability on scaled landscapes. Landscape Ecol. **8**, 213-227.

Van Horne, B. 1989. Spatial configuration of avian habitats. Acta XX Congressus Internationalis Ornithologici Vol. IV: 2313-2319.

Vuillod, P. 1994. Paysage visible et aménagement: modélisations cartographiques et test sur le Haut-Jura. Ph.D. dissertation, Université de Franche-Comté, France.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecol. **3**, 385-397.

GENERAL CONCLUSIONS

In this dissertation, I assessed the feasibility of using atlas data to develop habitat relationship models for seven woodpecker species in the Jura, France, and in Switzerland. Although modern technologies such as GIS and statistical packages greatly facilitate the construction of such models, the modeler should be aware of a number of potential limitations and pitfalls inherent to using atlas data.

Format of the habitat variables

One of the major differences between atlas data and point-based distribution data (such as nest sites) is that habitat data will be measured within a grid cell rather than at a specific point. Deciding whether to dichotomize continuous habitat variables is a key decision facing a modeler, and it is one that is likely to influence the output of the model. For example, I chose to extract land cover data as percentages, i.e., the percent of each atlas cell composed of given vegetation classes (Chapters 1 and 3). This decision was based on the notion that, because the atlas cells were quite large (100 ha and 575 ha), the likelihood of a woodpecker's presence in a cell with less than 10% of a given forest cover type may differ from that in a cell composed of 90% of that forest cover type. The limitation of this approach is that the results will fail to predict a presence in cells dominated by unsuitable cover types, although the species may be present in that cell because of a small patch of suitable habitat. Therefore, errors of omission due to swamping of a suitable land cover type by an unsuitable one are likely. An alternative approach would consist in dichotomizing the land cover variables, giving a code of 0 if the cover type is absent in a given cell, and a code of 1 if it is present, regardless of the area of the cell occupied by that cover type. In this case, errors of

omission will be minimized, but errors of commission will be more abundant because the species may be predicted in any cell where the cover type is present.

Modeling approach

The decision of whether to give preference to commission or to omission errors will depend on the goal of the study. When conservation of potential habitat is the priority, omission errors will be more detrimental than commission ones -- provided that funding is available to survey areas where the species is not currently present. If funding is limited, then conservation efforts should focus on areas of suitable habitat that are currently occupied by the species.

The ability to control which type of error is minimized depends on the modeling approach. Because the output of logistic regression is probabilistic, a cut-off is usually defined by the modeler, below which the probabilities are considered absences and above which they are considered presences. The choice of this cut-off will have important consequences on the type of error that is minimized. In Chapter 1, I selected cut-off values with the goal of maximizing both the number of presences and absences correctly classified. There are, however, a variety of ways to decide how to select the cut-off value, and the method used is likely to influence the type of prediction error that is minimized.

The modeler will also have to decide which criteria should be used to measure the classification success of the model. I chose to present Cohen's Kappa over (or in addition to) percent presence and absence correctly classified because the ratio of actual presences over absences differed from 0.5 for most of the woodpecker species (Chapters 1 and 3). A larger proportion of actual presences or absences may lead to high percent presences or absences correctly classified because of chance alone (Titus *et*

al. 1994). Because of the way it is computed (Appendix 2), Cohen's Kappa will be sensitive to the cut-off value. I did not look at the influence of different approaches to select the cut-off on the value of Kappa, so the methods I used in Chapters 1 and 3 may not have maximized Kappa. Looking at the relationship between cut-off values and Kappa values would make a valuable addition to the present study.

Availability of habitat data

These different factors (format of the habitat variables, selection of a cut-off value and of a measure of classification success) are within the control of the modeler. There are, however, other factors inherent to using atlas data that cannot be controlled, and that both modeler and users should be aware of. Atlas data provide an opportunity to develop models for relatively large areas (e.g., 5055 km^2 for the Jura, Chapter 1), but it will be possible to build meaningful models (in terms of species conservation and management) only if habitat data are available for the corresponding area. Although remote sensing technology has revolutionized the ease with which land cover information can be obtained over broad areas, it cannot capture fine-scale habitat features such as tree snags. The possibility of obtaining information other than land cover composition, though, is increasing, as better classification systems are developed. For example, satellite data may also be used to derive information about canopy closure and stand structure. I was unable to obtain fine-scale information, such as snag density or presence of specific micro-habitats, for the whole study area, and this resulted in lower classification success of some of the models (e.g., grey-headed woodpecker, Chapter 1). If the models are to be used for habitat management, the variables entering them should be those upon which management action can be taken. Such variables may not be available for a large study area, or if they are, their scale may not be compatible

with that of the atlas cell size (problem of data swamping).

Spatial autocorrelation

Another problem facing the user of atlas data is that of spatial autocorrelation, i.e., non-independence of the residuals. Spatial autocorrelation will occur if the presence of a species in an atlas cell is not independent of its presence in neighboring cells. Techniques are available to measure the amount of spatial autocorrelation in the data, such as the construction of semi-variograms (Appendix 1). If spatial autocorrelation is detected and deemed to be important enough to significantly affect the models two courses of action are available. First, neighboring or adjacent cells can be eliminated from analysis (e.g. Gates *et al.* 1993), or only one in every *n* cells can be kept (e.g. Pereira and Itami 1991). This solution will only be possible if the original atlas is composed of a large number of cells, so that eliminating some from analysis does not lead to sample sizes that are too small to ensure model stability. A second approach consists in explicitly modeling spatial autocorrelation (e.g. Smith 1994, Augustin *et al.* 1996). Because it will require that the actual distribution of the species be known, this approach is not recommended if the models are developed with the goal of predicting species' distributions outside of the original study area (unless of course distribution data are available). This can be a limitation if the models are intended for simulations, for example to compare alternative management scenarios on the potential distribution of species.

Grid positioning problems

Other limitations of atlas data will be more difficult to address. For example, grid positioning problems (Chapter 1) are inherent to working with gridded data, and

will always lead to at least some errors of commission or omission. Fielding and Haworth (1995) encountered such problems when modeling the distribution of golden eagles (*Aquila chrysaetos*) in several Scottish islands. When nests were located along the coast, cell composition was dominated by sea, and the models failed to predict a presence. I faced similar problems in Chapter 1. For example, the variable Open negatively entered models for the black woodpecker, so if a bird nested in a the corner of a large forest patch, but in a cell otherwise dominated by open habitat, the model would misclassify that cell. Little can be done to solve such grid positioning issues.

Scale issues

Scale is likely to be an important issue when developing models from atlas data. In Chapter 3, I varied the minimum mapping unit of the habitat map, and changed the cell size of the original atlas grid by grouping four, 100-ha cells into one, 400-ha cell. Both manipulations (MMU and cell size) influenced classification success, and the sign of this influence (positive or negative) was a function of the study area. For example, increasing the MMU lead to a lower classification success in Geneva Canton, but had little effect in the Orbe Valley. Because species-habitat associations are scale-dependent (Wiens 1989), and because atlas data come in a fixed, unique cell size, the modeler should be aware that the correlations between a given species' presence and habitat variables may be artifacts, or if they are biologically meaningful, that a model developed elsewhere with atlas data using a different cell size may lead to different results. Influence of atlas cell size is also likely to be species-specific, because different species use their environment at different scales. For example, the black woodpecker occupies larger territories than the great-spotted woodpecker; a single, 100-ha cell may comprise the territory of the latter, but not of the former.

Because scale is also a function of the geographic extent of the study area (Turner 1990), models are likely to be affected by the extent over which atlas censuses were conducted. Atlases usually correspond to administrative entities, which may bear little resemblance to natural range boundaries. The larger the area covered, the more habitat types it is likely to comprise (e.g., comparison of the Jura study area to the Orbe Valley or Geneva Canton study areas). Negative correlations were found between the black woodpecker's presence and deciduous forest types in the Jura (Chapter 1, Table 2); but had the study site been all of France instead of the Jura, the recent expansion of the species to lowland deciduous forests may have lead to positive correlations and to different models. As a consequence, models developed using data collected over a given area may not be applicable to smaller, or larger areas (e.g., Chapter 2).

Spatial variables

Another limitation of atlas data -- or, more generally, of data presented in a grid format as opposed to a point format -- is the inability to compute spatial variables. The spatial characteristics of landscape patches, in addition to their composition, may influence the distribution of wildlife species (Van Horne 1989). Several studies have suggested that patch attributes, such as size, may affect the distribution of bird species (e.g.Whitcomb *et al.* 1981, Lynch and Whigham 1984). Unfortunately, because each atlas cell is likely to overlap several patches, it is not possible to know which patch to consider. Computing mean patch size could lead to confusing results, because a cell with several, medium-sized patches could give the same outcome as one with one large patch and several small ones. In Chapter 3, differences in landscape configuration between the Orbe valley and Geneva Canton could not be accounted for because of this inability to compute patch statistics, and this in turn may have played a role in the poor

ability of models developed for the green woodpecker in one site to predict the distribution of this species in the other. I was also unable to verify whether the middle spotted woodpecker preferred patches larger than 30 ha, and avoided those smaller than 5 ha, as suggested by Cramp (1985).

Collection of the atlas data

Finally, the user of atlas data will have to pay particular attention to the way the data were collected. Did the authors make sure that sampling effort was equal for all cells, or is it possible that certain cells were surveyed in greater depth than others? Because atlas surveys often rely on a network of volunteers, cells most familiar to the survey crew, or cells easily accessible, may receive more attention than more isolated, difficult to reach, or topographically challenging cells. How long did the survey last? Data for the Jura Breeding Bird Atlas were collected over an eight-year period (Jovéniaux 1993). Although land cover characteristics did not change much over that period in the Jura, this may be different in other areas. If land cover characteristics undergo important changes between the time the atlas is started and the time surveys are over, models will be influenced by the date at which habitat variables are collected, and may fail to reflect true species-habitat associations. Using survey data collected over several years may lead to misleading results, especially if population density is highly variable. For example, when population densities are high, individuals are more likely to occupy lower quality habitats. Associations between species distribution and land cover types may lead the modeler to falsely conclude that poor quality habitats (and potential sinks) represent important types that deserve management and conservation attention.

The different problems outlined above are by no mean intended to discourage the use of atlas data in wildlife-habitat relationship models. Rather, they underscore the importance of exercising caution and scrutiny when developing and interpreting models, because decisions based on such models can have dramatic consequences for the long-term persistence of real, live organisms.

REFERENCES

Augustin, N. H., M. A. Mugglestone, and S. T. Buckland. 1996. An autologistic model for the spatial distribution of wildlife. J. Appl. Ecol. **33**, 339-347.

Cramp, S. 1985. The Birds of Western Palaearctic, Vol. IV. Oxford Univ. Press, Oxford.

Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conserv. Biol. **9**, 1466-1481.

Gates, S., D. W. Gibbons, P. C. Lack, and R. J. Fuller. 1993. Declining farmland bird species: modelling geographical patterns of abundance in Britain. Pages 153-177 *in* P. J. Edwards, R. M. May, and N. R. Webb, eds. Large-Scale Ecology and Conservation Biology. Blackwell Sci. Publ., London.

Jovéniaux, A. 1993. Atlas des Oiseaux Nicheurs du Jura. G.O.J. ed., Lons-le-Saunier.

Lynch, J. F., and D. F. Whigham. 1984. Effects of forest fragmentation on breeding bird communities in Maryland, U.S.A. Biol. Conserv. **28**, 287-324.

Pereira, J. M. C., and R. M. Itami. 1991. GIS-based habitat modeling using logistic regression: a study of the Mt. Graham red squirrel. Photogrammetric Engineering and Remote Sensing **57**, 1475-1486.

Smith, P. A. 1994. Autocorrelation in logistic regression modelling of species' distribution. Global Ecol. and Biogeography Letters **4**, 47-61.

Titus, K., J. A. Mosher, and B. K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: ecological applications. Am. Midl. Nat. **111**, 1-7.

Turner, M. G. 1990. Spatial and temporal analysis of landscape patterns. Landscape Ecol. **4**, 21-30.

Van Horne, B. 1989. Spatial configuration of avian habitats. Acta XX Congressus Internationalis Ornithologici Vol. IV: 2313-2319.

Whitcomb, R. F., C. S. Robbins, J. F. Lynch, B. L. Whitcomb, M. K. Klimjievicz, and D. Bystrak. 1981. Effects of forest fragmentation on avifauna of the eastern deciduous forest. Pages 123-205 *in* R. L. Burgess and D. M. Sharpe, eds. Forest Island Dynamics in Man-dominated Landscapes. Springer-Verlag, New York.

Wiens, J. A. 1989. The ecology of bird communities. Vol. 1. Foundations and patterns. Cambridge Univ. Press.

APPENDIX 1. MEASURING THE SPATIAL AUTOCORRELATION OF HABITAT VARIABLES IN THE JURA, FRANCE.

INTRODUCTION AND METHODS

Because of their presentation (in a grid format), data obtained from atlases may be prone to spatial autocorrelation, a phenomenon frequently encountered when a variable is mapped onto a geographic space (Sokal and Oden 1978). Most classical statistics assume independence of the observations, so statistical problems are to be expected; for example, spatial autocorrelation can lead to a poor model fit. To estimate the level of spatial autocorrelation in our data, I computed Moran's *I* and Geary's *c* indices for each habitat variable using the commands available in Arc/Info 7.0.3. (ESRI 1995). The closer *I* is to 0, and the closer *c* is to 1, the weaker the spatial autocorrelation (Goodchild 1986). The output of these two procedures, however, is a unique index for the entire map, because the indices are computed for adjacent cells only (ESRI 1995) and thus it is not possible to assess the influence of various lag distances on spatial autocorrelation . To explore spatial structure in the data, I conducted semivariance analysis as described by Marks and Aronson (1984) and applied by Fielding and Haworth (1995), a technique that permits the identification of ranges of scales where spatial autocorrelation is present (Bian and Walsh 1993). Because of the importance of the elevation gradient in the study area, I first assigned the 856 cells to one of three zones based on their elevation: less than 400 m, 400 m to 800 m, and higher than 800 m. I randomly selected 100 cells within each zone, then for each selected cell followed a 'transect' along one of the four cardinal directions. The direction of the transect was chosen randomly, as long as 10 adjacent cells were present

along it; transects that exited the study area before reaching this minimum length were not retained. For each of the 17 variables, and for all 13 vegetation types grouped together, I computed the Euclidean distance between the selected cell and the first 12 cells along the retained transect. The number twelve was retained to avoid directional bias (the Jura is longer along its NS axis; the minimum EW distance is about 30 km, i.e., 12 cells). Information on spatial autocorrelation is most crucial for the first few cells, so the few transects that had only 10 cells were retained. This selection process resulted in 3564 distances, corresponding to twelve 2.5-km lag classes (roughly the cell size). The mean squared distance d^2 was calculated for each lag class. The fractal dimension D was obtained from the equation 3 - (b/2), where b is the slope of the line of ln(d^2) against ln(lag distance), i.e., the slope of the semivariogram (Mark and Aronson 1984). A large D value (and a small b value) indicate a complex habitat surface, i.e., little spatial autocorrelation; the contrary (large b value, small D value) indicate strong spatial autocorrelation (Fielding and Haworth 1995). Semivariograms were visually checked for 'break points' indicating different domains of scale (Mark and Aronson 1984).

RESULTS AND DISCUSSION

Spatial autocorrelation between each cell and its neighbors was measured by Moran's I and Geary's c (Table 1). Moran's I varied from more than 0.7 for elevation, slope, mature deciduous and xeric forests, to less than 0.3 for scree forests and other cuts. Mirror-image results were obtained for Geary's c, with the lowest value for elevation (0.0292) and the highest for scree forests (0.7105). These values indicate

Table 1. Autocorrelation measures computed for 17 habitat variables in the Jura. France.

Habitat variable	Moran's *I*	Geary's *c*
Elevation	0.8717	0.0292
Slope	0.7348	0.1955
Edge density	0.4544	0.4872
Richness	0.4748	0.4725
Mature deciduous forests	0.7011	0.2593
Plain coppice	0.2735	0.5851
Xeric forests	0.7190	0.2432
Scree forests	0.2634	0.7105
Mountain deciduous forests	0.5305	0.4467
Beech forests	0.4608	0.5151
Fir forests	0.6437	0.3110
Spruce forests	0.6677	0.2133
Regeneration cuts	0.4659	0.5194
Other cuts	0.2659	0.7017
Patchy forests	0.3097	0.6483
Riparian & heterogeneous forests	0.4220	0.5632
Not forested	0.4973	0.4565

that, when only adjacent cells are considered, the level of spatial autocorrelation varies widely among variables, with topographic variables and mature deciduous and xeric forests exhibiting the highest levels of spatial dependency.

For all 18 variograms (17 variables used in model development, plus the multivariate sum of the 13 vegetation classes), breaks in the slope could be identified by visual inspection. A first break point was always evident, at lag distances ranging from 5 to 12.5 km (Table 2, Fig. 1). A second break point was sometimes present, although generally harder to locate, except for elevation and slope (lag = 17.5 km). The upper portion of the curve was difficult to analyze, often presenting a jagged pattern with a short periodicity, or sometimes reversing, thus reflecting negative spatial autocorrelation (Table 2). Because of the non-linear nature of most relationships, obtaining *D*-values for the whole curve using least-squares regression analysis would not have been appropriate. Instead, I computed *D* for the lower part of the curve (i.e., before the first break point). Fractal dimensions ranged from 2.39 for elevation and xeric forests, to 2.83 for edge density (Table 2). Scree forests exhibited the least spatial autocorrelation of the 12 forest classes ($D = 2.79$). When a multivariate distance was computed for the 132 vegetation classes, a relatively high autocorrelation ($D = 2.56$) was observed.

The different methods I used to measure spatial autocorrelation (Moran's *I*, Geary's *c*, fractal dimension) identified its presence in the data, and the results show that important information may be lost when only adjacent cells are used in computations (as in *I* and *c*). A good example is provided by slope. The I and c values (0.7348 and 0.1955, respectively) indicate strong spatial dependency, but this result is contradicted by a fairly high D value for the lower part of the variogram (2.77). This indicates that while spatial autocorrelation is present at a fine scale (i.e., for adjacent

Table 2. First break point and fractal dimension for the lower part and the whole curve of 18 habitat variable semi-variograms for the Jura, France.

Variables	Lower part	Break point (km)	Whole curve
Elevation	2.39	10	2.45
Slope	2.77	10	2.82
Richness	2.77	7.5	2.89
Edge density	2.83	7.5	2.91
Mature Deciduous	2.62	10	2.76
Plain Coppice	[3]	5	2.88
Xeric	2.39	12.5	2.65
Scree	2.79	7.5	2.98
Mountain deciduous	2.66	5	2.85
Beech	2.77	7.5	2.98
Fir	2.63	10	2.83
Spruce	2.54	5	2.73
Shelterwood cut	2.53	5	2.71
Other cuts	2.71	5	2.86
Patchy	2.77	5	2.98
Heterogeneous	2.65	5	2.85
Non forest	2.70	7.5	2.88
All vegetation	2.56	5	2.82

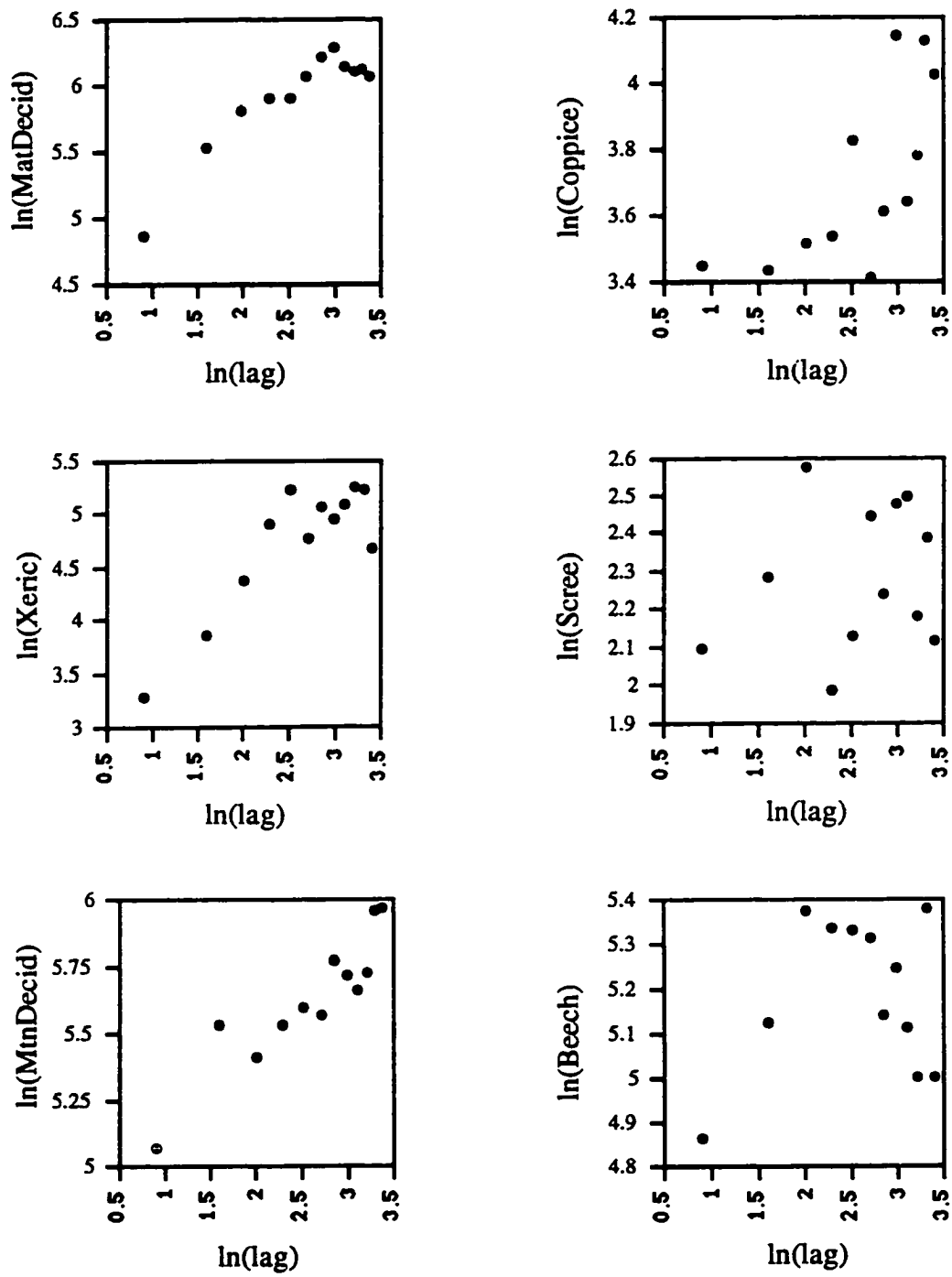


Figure 1. Semivariograms for 18 habitat variables in the Jura, France.

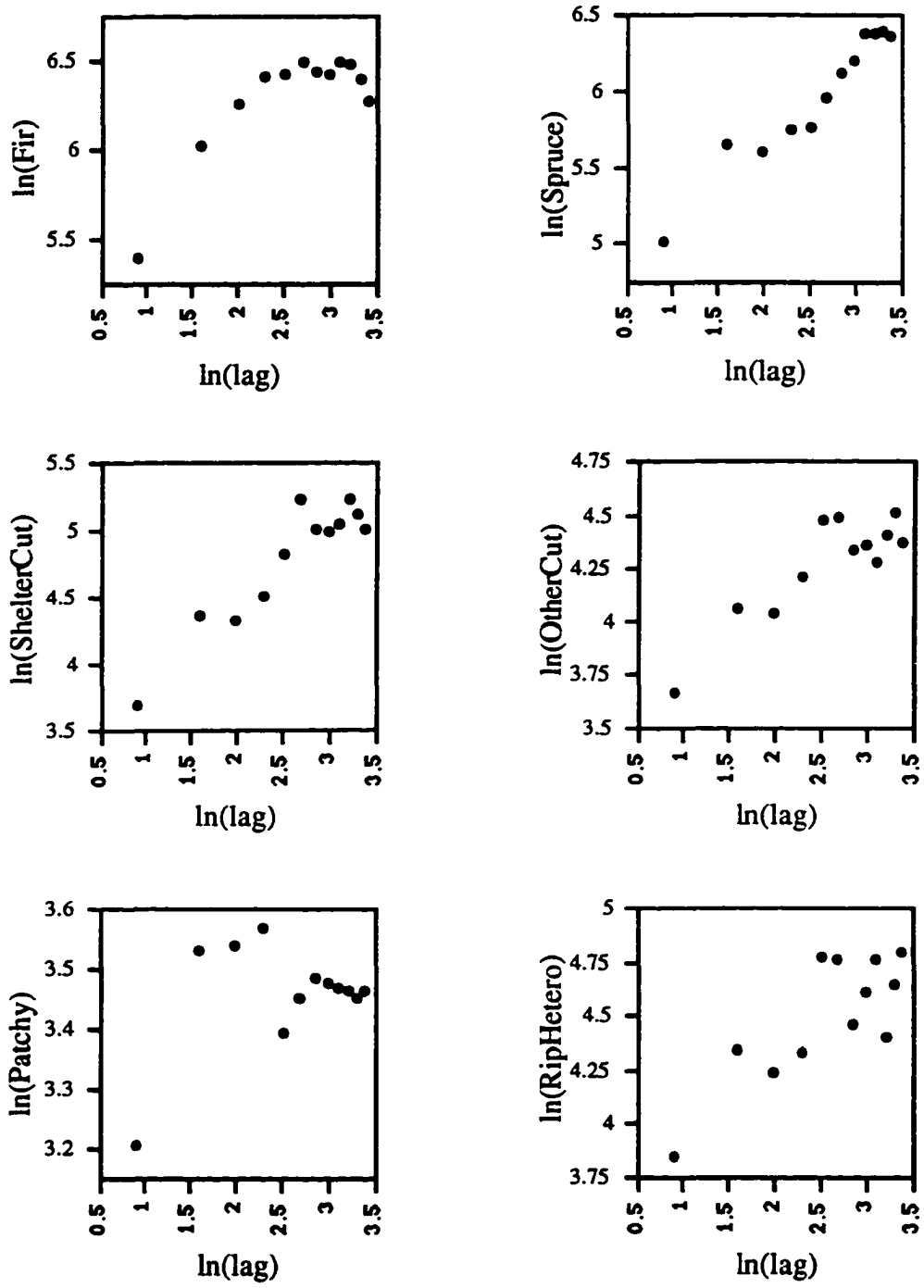


Figure 1 (continued). Semivariograms for 18 habitat variables in the Jura, France.

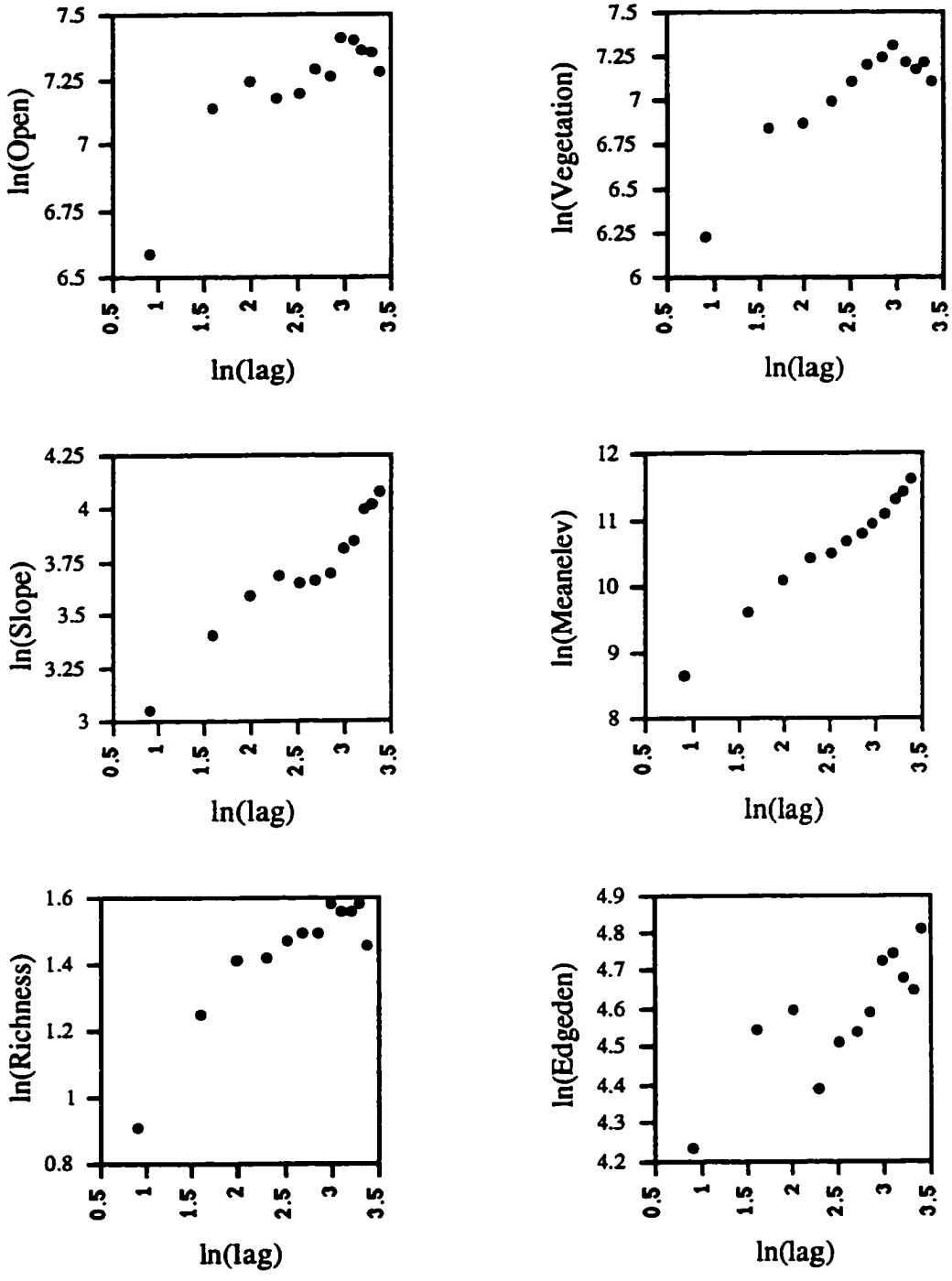


Figure 1 (continued). Semivariograms for 18 habitat variables in the Jura, France.

cells), it tends to decrease over broader areas (i.e., the 10-km lag before the first break point), simply because there are few slopes of constant value over such long planar distances. The variograms, then, may give a truer representation of spatial dependency in the data.

The lag distance of the first break point (from 5 to 12.5 km) reflects the topography of the study area, i.e., a succession of physiographic entities along a northeast-southwest altitudinal gradient. Mark and Aronson (1984) and Fielding and Haworth (1995) also noticed break-points at similar lag distances (5 km for both studies), even though they worked in very different landscapes. Except for the plains, which occupy the whole northern part of the Jura, physiographic entities in the Jura occur as long, narrow bands about 10 km wide (Fig. 1, Chapter 1). Break points, then, can either be transitions between physiographic entities, or reflect the spatial distribution of cover types, or both. For example, all the spruce forests are found in the Upper Jura entity, so the break point for this type probably corresponds to the Upper Jura/Second Plateau transition. On the other hand, spatial autocorrelation for mature deciduous forests is certainly strongly influenced by the fact that 40% of this forest type occur as a single patch, the Chaux forest. Similarly, xeric forests are restricted geographically to one area in the south-central part of the Jura, and exhibit the lowest *D*-value (the highest level of spatial autocorrelation) of the variables. Variables distributed in large, homogeneous patches are expected to show a high level of spatial autocorrelation (Sokal and Oden 1978). Conversely, this may explain why scree forests show spatial independence: this forest type is the rarest (0.52 %) and occurs in small, elongated patches scattered across the study area, usually at the junction between the Plains and the Little Mountain where scree slopes are most common. Man-related land uses, such as urban development or agriculture, can be expected to be unpatterned at landscape-

scales of analysis (O'Neill *et al.*1991). Although spatial autocorrelation is still present ($D = 2.71$), forest practices regrouped under 'other cuts' are scattered in the landscape independently of elevation. Shelterwood cuts, however, are much more clustered in the landscape (broad patches in the Moidons forest), and get a lower D-value ($D = 2.53$).

Plain coppice presents a special case. With an I value of 0.2735 and a c value of 0.5851, this forest type presents very low levels of spatial autocorrelation when only adjacent cells are considered. The corresponding variogram differs from all the other variograms, as it is the only one with a D value of 3 for the lower part of the curve - meaning that, at short lag distances, the variables actually exhibits negative spatial autocorrelation. The trend, however, reverses after 5-km lags and periodicity is observed, with phases of positive autocorrelation. The high D value may be explained by the general distribution pattern of plain coppice; it occurs in small patches scattered throughout the low-elevation part of the Jura, and rarely present in adjacent cells. I suspect, however, that the negative autocorrelation indicated by the variogram could be an artifact of the method used (only one out of eight adjacent cells was considered each time). Leduc *et al.* (1994) demonstrated that the size of the sampling unit, the relative position of the sampling grid or transect, and the orientation of the transect all affect the fractal dimension. Because of software limitations I computed D-values from transects running in the four cardinal directions only, so that any amount of spatial autocorrelation occurring diagonally was not picked up by the sampling scheme. As a result, spatial dependency is probably underestimated for certain cover types.

Fielding and Haworth (1995), working with 1km-cells (100 ha), found D-values ranging between 2.78 and 2.86 for 5-km lags, and concluded that the spatial dependency of their data was weak. However, they obtained only one value per

landscape by computing a multivariate fractal dimension. When performing a similar computation for all 13 vegetation types, I obtained a lower *D*-value than theirs (*D* = 2.56; Table 2), suggesting a higher level of spatial autocorrelation in our data. Working at a 575-ha scale masks small-scale patterns such as the succession of valleys and ridges typical of certain portions of the Jura, but uncovers larger geographic phenomena - the transition from plains to mountains. To drop Moran's I from 0.7 to 0.3 for elevation, Pereira and Itami (1992) chose to select only one out of seven cells for analysis. Gates *et al.* (1993), working with atlas data, use only non-adjacent cells. Removing samples will lower spatial autocorrelation, but it is not recommended because it requires valuable information to be discarded from analysis (Legendre 1993), and it reduces sample size. Spatial dependency, in my study, was so tied to the geomorphological structure of the Jura, that removing it would have entailed eliminating many cells from analysis. Small sample sizes usually lead to unstable LR models (Capen *et al.*1986). Another way to deal with spatial autocorrelation is to model it explicitly (Smith 1994, Augustin *et al.* 1996). I used the method described by Augustin *et al.* (1996) to predict the distribution of the grey-headed woodpecker and obtained a much tighter fit to the presence-absence data. However, because this approach requires that the distribution of the modeled species be known beforehand, the resulting models cannot be used for prediction in areas where this distribution is unknown (Smith 1994). I therefore chose not to remove the spatial dependency from my data, even though removing it may have lead to a better fit of the models.

REFERENCES

Augustin, N. H., M. A. Mugglestone, and S. T. Buckland. 1996. An autologistic model for the spatial distribution of wildlife. J. Appl. Ecol. **33**, 339-347.

Bian, L., and S. J. Walsh. 1993. Scale dependencies of vegetation and topography in a mountainous environment of Montana. Professional Geographer **45**, 1-11.

Capen, D. E., J. W. Fenwick, D. B. Inkley, and A. C. Boynton. 1986. Multivariate models of songbird habitat in New England forests. Pages 171-175 *in* J. Verner, M. L. Morrison, and C. J. Ralph, eds. Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates. Univ. of Wisconsin Press, Madison.

E.S.R.I. Arc/Info 7.0.3. Environmental Systems Research Institute, Redlands, CA.

Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conserv. Biol. **9**, 1466-1481.

Gates, S., D. W. Gibbons, P. C. Lack, and R. J. Fuller. 1993. Declining farmland bird species: modelling geographical patterns of abundance in Britain. Pages 153-177 *in* P. J. Edwards, R. M. May, and N. R. Webb, eds. Large-Scale Ecology and Conservation Biology. Blackwell Sci. Publ., London.

Goodchild, M. F. 1986. Spatial Autocorrelation. Catmog 47, Geobooks, Norwich.

Leduc, A., Y. T. Prairie, and Y. Bergeron. 1994. Fractal dimension estimates of a fragmented landscape: sources of variability. Landscape Ecol. **9**, 279-286.

Mark, D. M., and P. B. Aronson. 1984. Scale-dependent fractal dimensions of topographic surfaces: an empirical investigation, with applications in geomorphology and computer mapping. Mathematical Geol. **16**, 671-683.

O'Neill, R. V., R. H. Gardner, B. T. Milne, M. G. Turner, and B. Jackson. 1991. Heterogeneity and spatial hierarchies. Pages 85-96 *in* J. Kolasa and S.T.A. Pickett, eds. Ecological Heterogeneity. Ecological Studies 86, Springer-Verlag.

Pereira, J. M. C., and R. M. Itami. 1991. GIS-based habitat modeling using logistic regression: a study of the Mt. Graham red squirrel. Photogrammetric Engineering and Remote Sensing **57**, 1475-1486.

Smith, P. A. 1994. Autocorrelation in logistic regression modelling of species' distribution. Global Ecol. and Biogeography Letters **4**, 47-61.

Sokal, R. R., and N. L. Oden. 1978. Spatial autocorrelation in biology. 1. Methodology. Biol. J. of the Linnean Soc. **10**, 199-228.

APPENDIX 2. FORMULAS USED FOR THE COMPUTATION OF COHEN'S KAPPA (K) AND THE CORRESPONDING P-VALUE.

The example presented here uses data of the biological model of the black woodpecker (Chapter 1). Differences between the values in this Appendix and those listed in Table 6 (Chapter 1) are due to rounding errors. Formulas are from Titus *et al.* (1994)*.

	Predicted presences	Predicted absences	Total
True presences	98	22	120 (0.140)[a]
True absences	166	570	736 (0.860)[a]
Total	264 (0.308)[b]	592 (0.692)[b]	856

[a, b]: Letter superscript refer to calculations of row (a) and column (b) proportions.

P_o = the sum of the observed portion of agreement

P_o = (Σ matrix diagonal elements) / (Σ all elements) = (98 + 570) / 856 = 0.780

P_c = the sum of the chance expected proportion of agreement

a. Row proportions = (Σ columns for each row) / (Σ all elements)

b. Column proportions = (Σ rows for each column) / (Σ all elements)

c. P_c = Σ (row proportion X column proportion)

= (0.308 X 0.140) + (0.692 X 0.860) = 0.638

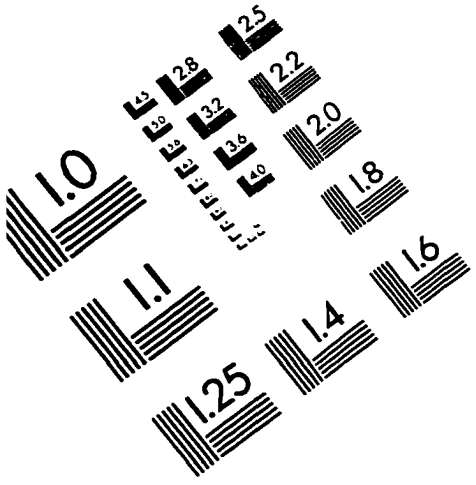
$K = (P_o - P_c) / (1 - P_c) = (0.780 - 0.638) / (1 - 0.638) = 0.392$

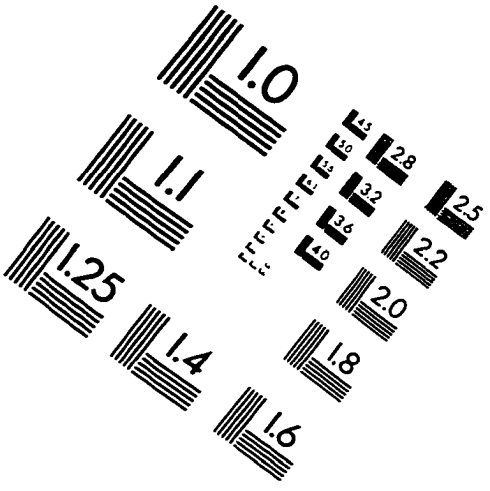
$$SE_k = \sqrt{\frac{Pc}{N\,(1-Pc)}} = \sqrt{\frac{0.638}{856\,(1-0.638)}} = 0.045$$

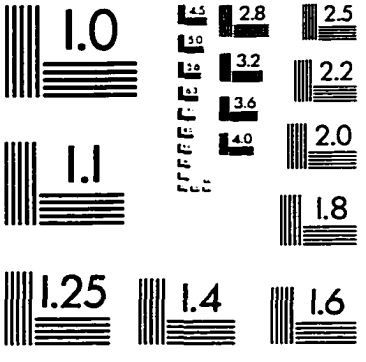
$Z = (K - 0) / SE_k = 0.392 / 0.045 = 8.711$, $P < 0.0001$

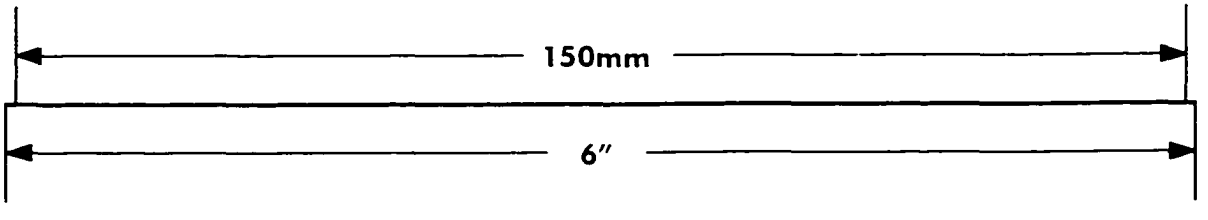
* Titus, K., J. A. Mosher, and B. K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: ecological applications. Am. Midl. Nat. **111**, 1-7.

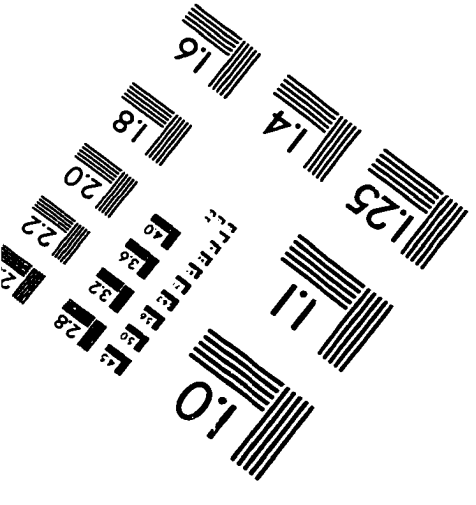
IMAGE EVALUATION TEST TARGET (QA-3)

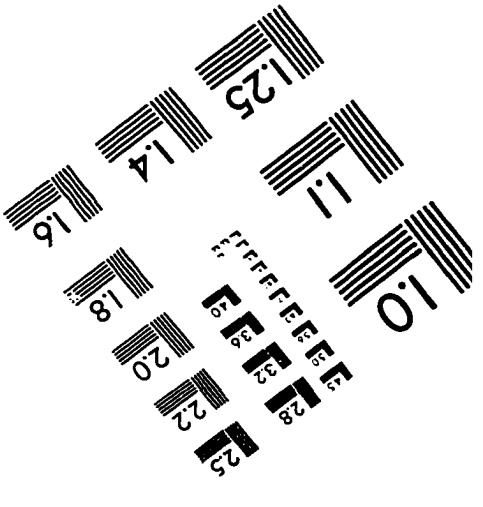












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989